

- 1, TITLE: Multimodal Model-Agnostic Meta-Learning via Task-Aware Modulation
<https://papers.nips.cc/paper/8296-multimodal-model-agnostic-meta-learning-via-task-aware-modulation>
AUTHORS: Risto Vuorio, Shao-Hua Sun, Hexiang Hu, Joseph J. Lim
HIGHLIGHT: In this paper, we augment MAML with the capability to identify the mode of tasks sampled from a multimodal task distribution and adapt quickly through gradient updates.

- 2, TITLE: ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks
<https://papers.nips.cc/paper/8297-vilbert-pretraining-task-agnostic-visiolinguistic-representations-for-vision-and-language-tasks>
AUTHORS: Jiasen Lu, Dhruv Batra, Devi Parikh, Stefan Lee
HIGHLIGHT: We present ViLBERT (short for Vision-and-Language BERT), a model for learning task-agnostic joint representations of image content and natural language.

- 3, TITLE: Stochastic Shared Embeddings: Data-driven Regularization of Embedding Layers
<https://papers.nips.cc/paper/8298-stochastic-shared-embeddings-data-driven-regularization-of-embedding-layers>
AUTHORS: Liwei Wu, Shuqing Li, Cho-Jui Hsieh, James L. Sharpnack
HIGHLIGHT: Alternatively, we propose stochastically shared embeddings (SSE), a data-driven approach to regularizing embedding layers, which stochastically transitions between embeddings during stochastic gradient descent (SGD).

- 4, TITLE: Unsupervised Scale-consistent Depth and Ego-motion Learning from Monocular Video
<https://papers.nips.cc/paper/8299-unsupervised-scale-consistent-depth-and-ego-motion-learning-from-monocular-video>
AUTHORS: Jiawang Bian, Zhichao Li, Naiyan Wang, Huangying Zhan, Chunhua Shen, Ming-Ming Cheng, Ian Reid
HIGHLIGHT: This paper tackles these challenges by proposing a geometry consistency loss for scale-consistent predictions and an induced self-discovered mask for handling moving objects and occlusions.

- 5, TITLE: Zero-shot Learning via Simultaneous Generating and Learning
<https://papers.nips.cc/paper/8300-zero-shot-learning-via-simultaneous-generating-and-learning>
AUTHORS: Hyeonwoo Yu, Beomhee Lee
HIGHLIGHT: Beyond exploiting relations between classes of seen and unseen, we present a deep generative model to provide the model with experience about both seen and unseen classes.

- 6, TITLE: Ask not what AI can do, but what AI should do: Towards a framework of task delegability
<https://papers.nips.cc/paper/8301-ask-not-what-ai-can-do-but-what-ai-should-do-towards-a-framework-of-task-delegability>
AUTHORS: Brian Lubars, Chenhao Tan
HIGHLIGHT: We approach this problem of task delegability from a human-centered perspective by developing a framework on human perception of task delegation to AI.

- 7, TITLE: Stand-Alone Self-Attention in Vision Models
<https://papers.nips.cc/paper/8302-stand-alone-self-attention-in-vision-models>
AUTHORS: Niki Parmar, Prajit Ramachandran, Ashish Vaswani, Irwan Bello, Anselm Levskaya, Jon Shlens
HIGHLIGHT: In developing and testing a pure self-attention vision model, we verify that self-attention can indeed be an effective stand-alone layer.

- 8, TITLE: High Fidelity Video Prediction with Large Stochastic Recurrent Neural Networks
<https://papers.nips.cc/paper/8303-high-fidelity-video-prediction-with-large-stochastic-recurrent-neural-networks>
AUTHORS: Ruben Villegas, Arkanath Pathak, Harini Kannan, Dumitru Erhan, Quoc V. Le, Honglak Lee
HIGHLIGHT: In this work, we question if such handcrafted architectures are necessary and instead propose a different approach: finding minimal inductive bias for video prediction while maximizing network capacity.

- 9, TITLE: Unsupervised learning of object structure and dynamics from videos
<https://papers.nips.cc/paper/8304-unsupervised-learning-of-object-structure-and-dynamics-from-videos>
AUTHORS: Matthias Minderer, Chen Sun, Ruben Villegas, Forrester Cole, Kevin P. Murphy, Honglak Lee
HIGHLIGHT: To address this challenge, we adopt a keypoint-based image representation and learn a stochastic dynamics model of the keypoints.

- 10, TITLE: GPipe: Efficient Training of Giant Neural Networks using Pipeline Parallelism
<https://papers.nips.cc/paper/8305-gpipe-efficient-training-of-giant-neural-networks-using-pipeline-parallelism>
AUTHORS: Yanping Huang, Youlong Cheng, Ankur Bapna, Orhan Firat, Dehao Chen, Mia Chen, HyoukJoong Lee, Jiquan Ngiam, Quoc V. Le, Yonghui Wu, zhifeng Chen
HIGHLIGHT: To address the need for efficient and task-independent model parallelism, we introduce TensorPipe, a pipeline parallelism library that allows scaling any network that can be expressed as a sequence of layers.

11, TITLE: Meta-Learning with Implicit Gradients
<https://papers.nips.cc/paper/8306-meta-learning-with-implicit-gradients>
AUTHORS: Aravind Rajeswaran, Chelsea Finn, Sham M. Kakade, Sergey Levine
HIGHLIGHT: By drawing upon implicit differentiation, we develop the implicit MAML algorithm, which depends only on the solution to the inner level optimization and not the path taken by the inner loop optimizer.

12, TITLE: Adversarial Examples Are Not Bugs, They Are Features
<https://papers.nips.cc/paper/8307-adversarial-examples-are-not-bugs-they-are-features>
AUTHORS: Andrew Ilyas, Shibani Santurkar, Dimitris Tsipras, Logan Engstrom, Brandon Tran, Aleksander Madry
HIGHLIGHT: We demonstrate that adversarial examples can be directly attributed to the presence of non-robust features: features (derived from patterns in the data distribution) that are highly predictive, yet brittle and (thus) incomprehensible to humans.

13, TITLE: Social-BiGAT: Multimodal Trajectory Forecasting using Bicycle-GAN and Graph Attention Networks
<https://papers.nips.cc/paper/8308-social-bigat-multimodal-trajectory-forecasting-using-bicycle-gan-and-graph-attention-networks>
AUTHORS: Vineet Kosaraju, Amir Sadeghian, Roberto Marten, Ian Reid, Hamid Rezaeifighi, Silvio Savarese
HIGHLIGHT: In this paper, we present Social-BiGAT, a graph-based generative adversarial network that generates realistic, multimodal trajectory predictions for multiple pedestrians in a scene.

14, TITLE: FreeAnchor: Learning to Match Anchors for Visual Object Detection
<https://papers.nips.cc/paper/8309-freeanchor-learning-to-match-anchors-for-visual-object-detection>
AUTHORS: Xiaosong Zhang, Fang Wan, Chang Liu, Rongrong Ji, Qixiang Ye
HIGHLIGHT: In this study, we propose a learning-to-match approach to break IoU restriction, allowing objects to match anchors in a flexible manner.

15, TITLE: Private Hypothesis Selection
<https://papers.nips.cc/paper/8310-private-hypothesis-selection>
AUTHORS: Mark Bun, Gautam Kamath, Thomas Steinke, Steven Z. Wu
HIGHLIGHT: We provide a differentially private algorithm for hypothesis selection.

16, TITLE: Differentially Private Algorithms for Learning Mixtures of Separated Gaussians
<https://papers.nips.cc/paper/8311-differentially-private-algorithms-for-learning-mixtures-of-separated-gaussians>
AUTHORS: Gautam Kamath, Or Sheffet, Vikrant Singhal, Jonathan Ullman
HIGHLIGHT: In this work, we give new algorithms for learning the parameters of a high-dimensional, well separated, Gaussian mixture model subject to the strong constraint of differential privacy.

17, TITLE: Average-Case Averages: Private Algorithms for Smooth Sensitivity and Mean Estimation
<https://papers.nips.cc/paper/8312-average-case-averages-private-algorithms-for-smooth-sensitivity-and-mean-estimation>
AUTHORS: Mark Bun, Thomas Steinke
HIGHLIGHT: We propose the trimmed mean estimator, which interpolates between the mean and the median, as a way of attaining much lower sensitivity on average while losing very little in terms of statistical accuracy.

18, TITLE: Multi-Resolution Weak Supervision for Sequential Data
<https://papers.nips.cc/paper/8313-multi-resolution-weak-supervision-for-sequential-data>
AUTHORS: Paroma Varma, Frederic Sala, Shiori Sagawa, Jason Fries, Daniel Fu, Saelig Khattar, Ashwini Ramamoorthy, Ke Xiao, Kayvon Fatahalian, James Priest, Christopher R?
HIGHLIGHT: We propose Dugong, the first framework to model multi-resolution weak supervision sources with complex correlations to assign probabilistic labels to training data.

19, TITLE: DeepUSPS: Deep Robust Unsupervised Saliency Prediction via Self-supervision
<https://papers.nips.cc/paper/8314-deepusps-deep-robust-unsupervised-saliency-prediction-via-self-supervision>
AUTHORS: Tam Nguyen, Maximilian Dax, Chaithanya Kumar Mummadi, Nhung Ngo, Thi Hoai Phuong Nguyen, Zhongyu Lou, Thomas Brox
HIGHLIGHT: In this work, we propose a two-stage mechanism for robust unsupervised object saliency prediction, where the first stage involves refinement of the noisy pseudo labels generated from different handcrafted methods.

20, TITLE: The Point Where Reality Meets Fantasy: Mixed Adversarial Generators for Image Splice Detection
<https://papers.nips.cc/paper/8315-the-point-where-reality-meets-fantasy-mixed-adversarial-generators-for-image-splice-detection>
AUTHORS: Vladimir V. Kniaz, Vladimir Knyaz, Fabio Remondino

HIGHLIGHT: In this paper, we propose a new framework for training of discriminative segmentation model via an adversarial process.

21, TITLE: You Only Propagate Once: Accelerating Adversarial Training via Maximal Principle
<https://papers.nips.cc/paper/8316-you-only-propagate-once-accelerating-adversarial-training-via-maximal-principle>
AUTHORS: Dinghui Zhang, Tianyuan Zhang, Yiping Lu, Zhanxing Zhu, Bin Dong
HIGHLIGHT: In this paper, we show that adversarial training can be cast as a discrete time differential game.

22, TITLE: Imitation Learning from Observations by Minimizing Inverse Dynamics Disagreement
<https://papers.nips.cc/paper/8317-imitation-learning-from-observations-by-minimizing-inverse-dynamics-disagreement>
AUTHORS: Chao Yang, Xiaojian Ma, Wenbing Huang, Fuchun Sun, Huaping Liu, Junzhou Huang, Chuang Gan
HIGHLIGHT: In this paper, we investigate LfO and its difference with LfD in both theoretical and practical perspectives.

23, TITLE: Asymptotic Guarantees for Learning Generative Models with the Sliced-Wasserstein Distance
<https://papers.nips.cc/paper/8318-asymptotic-guarantees-for-learning-generative-models-with-the-sliced-wasserstein-distance>
AUTHORS: Kimia Nadjahi, Alain Durmus, Umut Simsekli, Roland Badeau
HIGHLIGHT: In this study, we investigate the asymptotic properties of estimators that are obtained by minimizing SW.

24, TITLE: Generalized Sliced Wasserstein Distances
<https://papers.nips.cc/paper/8319-generalized-sliced-wasserstein-distances>
AUTHORS: Soheil Kolouri, Kimia Nadjahi, Umut Simsekli, Roland Badeau, Gustavo Rohde
HIGHLIGHT: In this paper, we first clarify the mathematical connection between the SW distance and the Radon transform. We then utilize the generalized Radon transform to define a new family of distances for probability measures, which we call generalized sliced-Wasserstein (GSW) distances.

25, TITLE: First Exit Time Analysis of Stochastic Gradient Descent Under Heavy-Tailed Gradient Noise
<https://papers.nips.cc/paper/8320-first-exit-time-analysis-of-stochastic-gradient-descent-under-heavy-tailed-gradient-noise>
AUTHORS: Thanh Huy Nguyen, Umut Simsekli, Mert Gurbuzbalaban, Gaël RICHARD
HIGHLIGHT: In this study, we provide formal theoretical analysis where we derive explicit conditions for the step-size such that the metastability behavior of the discrete-time system is similar to its continuous-time limit.

26, TITLE: Blind Super-Resolution Kernel Estimation using an Internal-GAN
<https://papers.nips.cc/paper/8321-blind-super-resolution-kernel-estimation-using-an-internal-gan>
AUTHORS: Sefi Bell-Kligler, Assaf Shocher, Michal Irani
HIGHLIGHT: In this paper we show how this powerful cross-scale recurrence property can be realized using Deep Internal Learning.

27, TITLE: Noise-tolerant fair classification
<https://papers.nips.cc/paper/8322-noise-tolerant-fair-classification>
AUTHORS: Alex Lamy, Ziyuan Zhong
HIGHLIGHT: In this paper, we answer the question in the affirmative: we show that if one measures fairness using the mean-difference score, and sensitive features are subject to noise from the mutually contaminated learning model, then owing to a simple identity we only need to change the desired fairness-tolerance.

28, TITLE: Generalization in Generative Adversarial Networks: A Novel Perspective from Privacy Protection
<https://papers.nips.cc/paper/8323-generalization-in-generative-adversarial-networks-a-novel-perspective-from-privacy-protection>
AUTHORS: Bingzhe Wu, Shiwang Zhao, Chaochao Chen, Haoyang Xu, Li Wang, Xiaolu Zhang, Guangyu Sun, Jun Zhou
HIGHLIGHT: In this paper, we aim to understand the generalization properties of generative adversarial networks (GANs) from a new perspective of privacy protection.

29, TITLE: Joint-task Self-supervised Learning for Temporal Correspondence
<https://papers.nips.cc/paper/8324-joint-task-self-supervised-learning-for-temporal-correspondence>
AUTHORS: Xueting Li, Sifei Liu, Shalini De Mello, Xiaolong Wang, Jan Kautz, Ming-Hsuan Yang
HIGHLIGHT: This paper proposes to learn reliable dense correspondence from videos in a self-supervised manner.

30, TITLE: Provable Gradient Variance Guarantees for Black-Box Variational Inference
<https://papers.nips.cc/paper/8325-provable-gradient-variance-guarantees-for-black-box-variational-inference>
AUTHORS: Justin Domke
HIGHLIGHT: Recent variational inference methods use stochastic gradient estimators whose variance is not well understood.

- 31, TITLE: Divide and Couple: Using Monte Carlo Variational Objectives for Posterior Approximation
<https://papers.nips.cc/paper/8326-divide-and-couple-using-monte-carlo-variational-objectives-for-posterior-approximation>
AUTHORS: Justin Domke, Daniel R. Sheldon
HIGHLIGHT: This paper gives bounds for the common “reparameterization” estimators when the target is smooth and the variational family is a location-scale distribution.
- 32, TITLE: Experience Replay for Continual Learning
<https://papers.nips.cc/paper/8327-experience-replay-for-continual-learning>
AUTHORS: David Rolnick, Arun Ahuja, Jonathan Schwarz, Timothy Lillicrap, Gregory Wayne
HIGHLIGHT: Here, we introduce CLEAR, a replay-based method that greatly reduces catastrophic forgetting in multi-task reinforcement learning.
- 33, TITLE: Deep ReLU Networks Have Surprisingly Few Activation Patterns
<https://papers.nips.cc/paper/8328-deep-relu-networks-have-surprisingly-few-activation-patterns>
AUTHORS: Boris Hanin, David Rolnick
HIGHLIGHT: In this paper, we show that the average number of activation patterns for ReLU networks at initialization is bounded by the total number of neurons raised to the input dimension.
- 34, TITLE: Chasing Ghosts: Instruction Following as Bayesian State Tracking
<https://papers.nips.cc/paper/8329-chasing-ghosts-instruction-following-as-bayesian-state-tracking>
AUTHORS: Peter Anderson, Ayush Srivastava, Devi Parikh, Dhruv Batra, Stefan Lee
HIGHLIGHT: Based on this intuition, we formulate the problem of finding the goal location in Vision-and-Language Navigation (VLN) within the framework of Bayesian state tracking - learning observation and motion models conditioned on these expectable events.
- 35, TITLE: Block Coordinate Regularization by Denoising
<https://papers.nips.cc/paper/8330-block-coordinate-regularization-by-denoising>
AUTHORS: Yu Sun, Jiaming Liu, Ulugbek Kamilov
HIGHLIGHT: In this work, we develop a new block coordinate RED algorithm that decomposes a large-scale estimation problem into a sequence of updates over a small subset of the unknown variables.
- 36, TITLE: Reducing Noise in GAN Training with Variance Reduced Extragradient
<https://papers.nips.cc/paper/8331-reducing-noise-in-gan-training-with-variance-reduced-extragradient>
AUTHORS: Tatjana Chavdarova, Gauthier Gidel, François Fleuret, Simon Lacoste-Julien
HIGHLIGHT: We address this issue with a novel stochastic variance-reduced extragradient (SVRE) optimization algorithm, which for a large class of games improves upon the previous convergence rates proposed in the literature.
- 37, TITLE: Learning Erdos-Renyi Random Graphs via Edge Detecting Queries
<https://papers.nips.cc/paper/8332-learning-erdos-renyi-random-graphs-via-edge-detecting-queries>
AUTHORS: Zihan Li, Matthias Fresacher, Jonathan Scarlett
HIGHLIGHT: In this paper, we consider the problem of learning an unknown graph via queries on groups of nodes, with the result indicating whether or not at least one edge is present among those nodes.
- 38, TITLE: A Primal-Dual link between GANs and Autoencoders
<https://papers.nips.cc/paper/8333-a-primal-dual-link-between-gans-and-autoencoders>
AUTHORS: Hisham Husain, Richard Nock, Robert C. Williamson
HIGHLIGHT: In this work, we study the GAN and WAE models and make two main discoveries.
- 39, TITLE: muSSP: Efficient Min-cost Flow Algorithm for Multi-object Tracking
<https://papers.nips.cc/paper/8334-mussp-efficient-min-cost-flow-algorithm-for-multi-object-tracking>
AUTHORS: Congchao Wang, Yizhi Wang, Yinxue Wang, Chiung-Ting Wu, Guoqiang Yu
HIGHLIGHT: In this paper, by exploiting the special structures and properties of the graphs formulated in MOT problems, we develop an efficient min-cost flow algorithm, namely, minimum-update Successive Shortest Path (muSSP).
- 40, TITLE: Category Anchor-Guided Unsupervised Domain Adaptation for Semantic Segmentation
<https://papers.nips.cc/paper/8335-category-anchor-guided-unsupervised-domain-adaptation-for-semantic-segmentation>
AUTHORS: Qiming ZHANG, Jing Zhang, Wei Liu, Dacheng Tao
HIGHLIGHT: In this paper, we propose a novel category anchor-guided (CAG) UDA model for semantic segmentation, which explicitly enforces category-aware feature alignment to learn shared discriminative features and classifiers simultaneously.

- 41, TITLE: Invert to Learn to Invert
<https://papers.nips.cc/paper/8336-invert-to-learn-to-invert>
AUTHORS: Patrick Putzky, Max Welling
HIGHLIGHT: In this work, we propose an iterative inverse model with constant memory that relies on invertible networks to avoid storing intermediate activations.
- 42, TITLE: Equitable Stable Matchings in Quadratic Time
<https://papers.nips.cc/paper/8337-equitable-stable-matchings-in-quadratic-time>
AUTHORS: Nikolaos Tziavelis, Ioannis Giannakopoulos, Katerina Doka, Nectarios Koziris, Panagiotis Karras
HIGHLIGHT: In this paper, we propose an alternative that is computationally simpler and achieves high equity too.
- 43, TITLE: Zero-Shot Semantic Segmentation
<https://papers.nips.cc/paper/8338-zero-shot-semantic-segmentation>
AUTHORS: Maxime Bucher, Tuan-Hung VU, Matthieu Cord, Patrick P?rez
HIGHLIGHT: In this paper, we introduce the new task of zero-shot semantic segmentation: learning pixel-wise classifiers for never-seen object categories with zero training examples.
- 44, TITLE: Metric Learning for Adversarial Robustness
<https://papers.nips.cc/paper/8339-metric-learning-for-adversarial-robustness>
AUTHORS: Chengzhi Mao, Ziyuan Zhong, Junfeng Yang, Carl Vondrick, Baishakhi Ray
HIGHLIGHT: Motivated by this observation, we propose to regularize the representation space under attack with metric learning to produce more robust classifiers.
- 45, TITLE: DISN: Deep Implicit Surface Network for High-quality Single-view 3D Reconstruction
<https://papers.nips.cc/paper/8340-dsn-deep-implicit-surface-network-for-high-quality-single-view-3d-reconstruction>
AUTHORS: Qiangeng Xu, Weiyue Wang, Duygu Ceylan, Radomir Mech, Ulrich Neumann
HIGHLIGHT: In this paper, we present DISN, a Deep Implicit Surface Network which can generate a high-quality detail-rich 3D mesh from a 2D image by predicting the underlying signed distance fields.
- 46, TITLE: Batched Multi-armed Bandits Problem
<https://papers.nips.cc/paper/8341-batched-multi-armed-bandits-problem>
AUTHORS: Zijun Gao, Yanjun Han, Zhimei Ren, Zhengqing Zhou
HIGHLIGHT: In this paper, we study the multi-armed bandit problem in the batched setting where the employed policy must split data into a small number of batches.
- 47, TITLE: vGraph: A Generative Model for Joint Community Detection and Node Representation Learning
<https://papers.nips.cc/paper/8342-vgraph-a-generative-model-for-joint-community-detection-and-node-representation-learning>
AUTHORS: Fan-Yun Sun, Meng Qu, Jordan Hoffmann, Chin-Wei Huang, Jian Tang
HIGHLIGHT: We propose a probabilistic generative model called vGraph to learn community membership and node representation collaboratively.
- 48, TITLE: Differentially Private Bayesian Linear Regression
<https://papers.nips.cc/paper/8343-differentially-private-bayesian-linear-regression>
AUTHORS: Garrett Bernstein, Daniel R. Sheldon
HIGHLIGHT: We investigate the problem of Bayesian linear regression, with the goal of computing posterior distributions that correctly quantify uncertainty given privately released statistics.
- 49, TITLE: Semantic Conditioned Dynamic Modulation for Temporal Sentence Grounding in Videos
<https://papers.nips.cc/paper/8344-semantic-conditioned-dynamic-modulation-for-temporal-sentence-grounding-in-videos>
AUTHORS: Yitian Yuan, Lin Ma, Jingwen Wang, Wei Liu, Wenwu Zhu
HIGHLIGHT: In this paper, we propose a novel semantic conditioned dynamic modulation (SCDM) mechanism, which relies on the sentence semantics to modulate the temporal convolution operations for better correlating and composing the sentence related video contents over time.
- 50, TITLE: AGEM: Solving Linear Inverse Problems via Deep Priors and Sampling
<https://papers.nips.cc/paper/8345-agem-solving-linear-inverse-problems-via-deep-priors-and-sampling>
AUTHORS: Bichuan Guo, Yuxing Han, Jiangtao Wen
HIGHLIGHT: In this paper we propose to use a denoising autoencoder (DAE) prior to simultaneously solve a linear inverse problem and estimate its noise parameter.

- 51, TITLE: CPM-Nets: Cross Partial Multi-View Networks
<https://papers.nips.cc/paper/8346-cpm-nets-cross-partial-multi-view-networks>
AUTHORS: Changqing Zhang, Zongbo Han, yajie cui, Huazhu Fu, Joey Tianyi Zhou, Qinghua Hu
HIGHLIGHT: To address the challenge, we propose a novel framework termed Cross Partial Multi-View Networks (CPM-Nets).
- 52, TITLE: Learning to Predict Layout-to-image Conditional Convolutions for Semantic Image Synthesis
<https://papers.nips.cc/paper/8347-learning-to-predict-layout-to-image-conditional-convolutions-for-semantic-image-synthesis>
AUTHORS: Xihui Liu, Guojun Yin, Jing Shao, Xiaogang Wang, hongsheng Li
HIGHLIGHT: In order to better exploit the semantic layout for the image generator, we propose to predict convolutional kernels conditioned on the semantic label map to generate the intermediate feature maps from the noise maps and eventually generate the images.
- 53, TITLE: Staying up to Date with Online Content Changes Using Reinforcement Learning for Scheduling
<https://papers.nips.cc/paper/8348-staying-up-to-date-with-online-content-changes-using-reinforcement-learning-for-scheduling>
AUTHORS: Andrey Kolobov, Yuval Peres, Cheng Lu, Eric J. Horvitz
HIGHLIGHT: We propose a novel optimization objective for this setting that has several practically desirable properties, and efficient algorithms for it with optimality guarantees even in the face of mixed content change observability and initially unknown change model parameters.
- 54, TITLE: SysCD: A System-Aware Parallel Coordinate Descent Algorithm
<https://papers.nips.cc/paper/8349-syscd-a-system-aware-parallel-coordinate-descent-algorithm>
AUTHORS: Nikolaos Ioannou, Celestine Mendler-D'ner, Thomas Parnell
HIGHLIGHT: In this paper we propose a novel parallel stochastic coordinate descent (SCD) algorithm with convergence guarantees that exhibits strong scalability.
- 55, TITLE: Importance Weighted Hierarchical Variational Inference
<https://papers.nips.cc/paper/8350-importance-weighted-hierarchical-variational-inference>
AUTHORS: Artem Sobolev, Dmitry P. Vetrov
HIGHLIGHT: To overcome this roadblock, we introduce a new family of variational upper bounds on a marginal log-density in the case of hierarchical models (also known as latent variable models).
- 56, TITLE: RSN: Randomized Subspace Newton
<https://papers.nips.cc/paper/8351-rsn-randomized-subspace-newton>
AUTHORS: Robert Gower, Dmitry Koralev, Felix Lieder, Peter Richtarik
HIGHLIGHT: We develop a randomized Newton method capable of solving learning problems with huge dimensional feature spaces, which is a common setting in applications such as medical imaging, genomics and seismology.
- 57, TITLE: Trust Region-Guided Proximal Policy Optimization
<https://papers.nips.cc/paper/8352-trust-region-guided-proximal-policy-optimization>
AUTHORS: Yuhui Wang, Hao He, Xiaoyang Tan, Yaozhong Gan
HIGHLIGHT: In this paper, we give an in-depth analysis on the exploration behavior of PPO, and show that PPO is prone to suffer from the risk of lack of exploration especially under the case of bad initialization, which may lead to the failure of training or being trapped in bad local optima.
- 58, TITLE: Adversarial Self-Defense for Cycle-Consistent GANs
<https://papers.nips.cc/paper/8353-adversarial-self-defense-for-cycle-consistent-gans>
AUTHORS: Dina Bashkurova, Ben Usman, Kate Saenko
HIGHLIGHT: In this paper, we show how such self-attacking behavior of unsupervised translation methods affects their performance and provide two defense techniques.
- 59, TITLE: Towards closing the gap between the theory and practice of SVRG
<https://papers.nips.cc/paper/8354-towards-closing-the-gap-between-the-theory-and-practice-of-svrg>
AUTHORS: Othmane Sebbouh, Nidham Gazagnadou, Samy Jelassi, Francis Bach, Robert Gower
HIGHLIGHT: Our first contribution is that we take several steps towards closing this gap.
- 60, TITLE: Uniform Error Bounds for Gaussian Process Regression with Application to Safe Control
<https://papers.nips.cc/paper/8355-uniform-error-bounds-for-gaussian-process-regression-with-application-to-safe-control>
AUTHORS: Armin Lederer, Jonas Umlauf, Sandra Hirche

HIGHLIGHT: In this paper, we employ the Gaussian process distribution and continuity arguments to derive a novel uniform error bound under weaker assumptions.

61, **TITLE:** ETNet: Error Transition Network for Arbitrary Style Transfer
<https://papers.nips.cc/paper/8356-etnet-error-transition-network-for-arbitrary-style-transfer>
AUTHORS: Chunjin Song, Zhijie Wu, Yang Zhou, Minglun Gong, Hui Huang
HIGHLIGHT: Inspired by the works on error-correction, instead, we propose a self-correcting model to predict what is wrong with the current stylization and refine it accordingly in an iterative manner.

62, **TITLE:** No Pressure! Addressing the Problem of Local Minima in Manifold Learning Algorithms
<https://papers.nips.cc/paper/8357-no-pressure-addressing-the-problem-of-local-minima-in-manifold-learning-algorithms>
AUTHORS: Max Vladymyrov
HIGHLIGHT: We propose a natural extension to several manifold learning methods aimed at identifying pressured points, i.e. points stuck in the poor local minima and have poor embedding quality.

63, **TITLE:** Deep Equilibrium Models
<https://papers.nips.cc/paper/8358-deep-equilibrium-models>
AUTHORS: Shaojie Bai, J. Zico Kolter, Vladlen Koltun
HIGHLIGHT: We present a new approach to modeling sequential data: the deep equilibrium model (DEQ).

64, **TITLE:** Saccader: Improving Accuracy of Hard Attention Models for Vision
<https://papers.nips.cc/paper/8359-saccader-improving-accuracy-of-hard-attention-models-for-vision>
AUTHORS: Gamaleldin Elsayed, Simon Kornblith, Quoc V. Le
HIGHLIGHT: Here, we propose a novel hard attention model, which we term Saccader.

65, **TITLE:** Multiway clustering via tensor block models
<https://papers.nips.cc/paper/8360-multiway-clustering-via-tensor-block-models>
AUTHORS: Miaoyan Wang, Yuchen Zeng
HIGHLIGHT: We propose a tensor block model, develop a unified least-square estimation, and obtain the theoretical accuracy guarantees for multiway clustering.

66, **TITLE:** Regret Minimization for Reinforcement Learning with Vectorial Feedback and Complex Objectives
<https://papers.nips.cc/paper/8361-regret-minimization-for-reinforcement-learning-with-vectorial-feedback-and-complex-objectives>
AUTHORS: Wang Chi Cheung
HIGHLIGHT: We propose a no-regret algorithm based on the Frank-Wolfe algorithm (Frank and Wolfe 1956), UCRL2 (Jaksch et al. 2010), as well as a crucial and novel gradient threshold procedure.

67, **TITLE:** NAT: Neural Architecture Transformer for Accurate and Compact Architectures
<https://papers.nips.cc/paper/8362-nat-neural-architecture-transformer-for-accurate-and-compact-architectures>
AUTHORS: Yong Guo, Yin Zheng, Mingkui Tan, Qi Chen, Jian Chen, Peilin Zhao, Junzhou Huang
HIGHLIGHT: To make the problem feasible, we cast the optimization problem into a Markov decision process (MDP) and seek to learn a Neural Architecture Transformer (NAT) to replace the redundant operations with the more computationally efficient ones (e.g., skip connection or directly removing the connection).

68, **TITLE:** Selecting Optimal Decisions via Distributionally Robust Nearest-Neighbor Regression
<https://papers.nips.cc/paper/8363-selecting-optimal-decisions-via-distributionally-robust-nearest-neighbor-regression>
AUTHORS: Ruidi Chen, Ioannis Paschalidis
HIGHLIGHT: This paper develops a prediction-based prescriptive model for optimal decision making that (i) predicts the outcome under each action using a robust nonlinear model, and (ii) adopts a randomized prescriptive policy determined by the predicted outcomes.

69, **TITLE:** Network Pruning via Transformable Architecture Search
<https://papers.nips.cc/paper/8364-network-pruning-via-transformable-architecture-search>
AUTHORS: Xuanyi Dong, Yi Yang
HIGHLIGHT: To break the structure limitation of the pruned networks, we propose to apply neural architecture search to search directly for a network with flexible channel and layer sizes.

70, **TITLE:** Differentiable Cloth Simulation for Inverse Problems
<https://papers.nips.cc/paper/8365-differentiable-cloth-simulation-for-inverse-problems>
AUTHORS: Junbang Liang, Ming Lin, Vladlen Koltun

HIGHLIGHT: We propose a differentiable cloth simulator that can be embedded as a layer in deep neural networks.

71, TITLE: Poisson-Randomized Gamma Dynamical Systems
<https://papers.nips.cc/paper/8366-poisson-randomized-gamma-dynamical-systems>
AUTHORS: Aaron Schein, Scott Linderman, Mingyuan Zhou, David Blei, Hanna Wallach
HIGHLIGHT: This paper presents the Poisson-randomized gamma dynamical system (PRGDS), a model for sequentially observed count tensors that encodes a strong inductive bias toward sparsity and burstiness.

72, TITLE: Volumetric Correspondence Networks for Optical Flow
<https://papers.nips.cc/paper/8367-volumetric-correspondence-networks-for-optical-flow>
AUTHORS: Gengshan Yang, Deva Ramanan
HIGHLIGHT: Instead, we introduce several simple modifications that dramatically simplify the use of volumetric layers - (1) volumetric encoder-decoder architectures that efficiently capture large receptive fields, (2) multi-channel cost volumes that capture multi-dimensional notions of pixel similarities, and finally, (3) separable volumetric filtering that significantly reduces computation and parameters while preserving accuracy.

73, TITLE: Learning Conditional Deformable Templates with Convolutional Networks
<https://papers.nips.cc/paper/8368-learning-conditional-deformable-templates-with-convolutional-networks>
AUTHORS: Adrian Dalca, Marianne Rakic, John Guttag, Mert Sabuncu
HIGHLIGHT: In this work, we present a probabilistic model and efficient learning strategy that yields either universal or conditional templates, jointly with a neural network that provides efficient alignment of the images to these templates.

74, TITLE: Fast Low-rank Metric Learning for Large-scale and High-dimensional Data
<https://papers.nips.cc/paper/8369-fast-low-rank-metric-learning-for-large-scale-and-high-dimensional-data>
AUTHORS: Han Liu, Zhizhong Han, Yu-Shen Liu, Ming Gu
HIGHLIGHT: To address this issue, we present a novel fast low-rank metric learning (FLRML) method.

75, TITLE: Efficient Symmetric Norm Regression via Linear Sketching
<https://papers.nips.cc/paper/8370-efficient-symmetric-norm-regression-via-linear-sketching>
AUTHORS: Zhao Song, Ruosong Wang, Lin Yang, Hongyang Zhang, Peilin Zhong
HIGHLIGHT: We provide efficient algorithms for overconstrained linear regression problems with size $n \times d$ when the loss function is a symmetric norm (a norm invariant under sign-flips and coordinate-permutations).

76, TITLE: RUBi: Reducing Unimodal Biases for Visual Question Answering
<https://papers.nips.cc/paper/8371-rubi-reducing-unimodal-biases-for-visual-question-answering>
AUTHORS: Remi Cadene, Corentin Dancette, Hedi Ben younes, Matthieu Cord, Devi Parikh
HIGHLIGHT: We propose RUBi, a new learning strategy to reduce biases in any VQA model.

77, TITLE: Why Can't I Dance in the Mall? Learning to Mitigate Scene Bias in Action Recognition
<https://papers.nips.cc/paper/8372-why-cant-i-dance-in-the-mall-learning-to-mitigate-scene-bias-in-action-recognition>
AUTHORS: Jinwoo Choi, Chen Gao, Joseph C. E. Messou, Jia-Bin Huang
HIGHLIGHT: In this paper, we propose to mitigate scene bias for video representation learning.

78, TITLE: NeurVPS: Neural Vanishing Point Scanning via Conic Convolution
<https://papers.nips.cc/paper/8373-neurvps-neural-vanishing-point-scanning-via-conic-convolution>
AUTHORS: Yichao Zhou, Haozhi Qi, Jingwei Huang, Yi Ma
HIGHLIGHT: In this work, we identify a canonical conic space in which the neural network can effectively compute the global geometric information of vanishing points locally, and we propose a novel operator named conic convolution that can be implemented as regular convolutions in this space.

79, TITLE: DATA: Differentiable ArchiTecture Approximation
<https://papers.nips.cc/paper/8374-data-differentiable-architecture-approximation>
AUTHORS: Jianlong Chang, ximbang zhang, Yiwen Guo, GAOFENG MENG, SHIMING XIANG, Chunhong Pan
HIGHLIGHT: To bridge this gap, we develop Differentiable ArchiTecture Approximation (DATA) with an Ensemble Gumbel-Softmax (EGS) estimator to automatically approximate architectures during searching and validating in a differentiable manner.

80, TITLE: Learn, Imagine and Create: Text-to-Image Generation from Prior Knowledge
<https://papers.nips.cc/paper/8375-learn-imagine-and-create-text-to-image-generation-from-prior-knowledge>
AUTHORS: Tingting Qiao, Jing Zhang, Duanqing Xu, Dacheng Tao

HIGHLIGHT: In this paper, and inspired by this process, we propose a novel text-to-image method called LeicaGAN to combine the above three phases in a unified framework.

81, **TITLE:** Memory-oriented Decoder for Light Field Salient Object Detection
<https://papers.nips.cc/paper/8376-memory-oriented-decoder-for-light-field-salient-object-detection>
AUTHORS: Miao Zhang, Jingjing Li, JI WEI, Yongri Piao, Huchuan Lu
HIGHLIGHT: In this paper, we present a deep-learning-based method where a novel memory-oriented decoder is tailored for light field saliency detection.

82, **TITLE:** Multi-label Co-regularization for Semi-supervised Facial Action Unit Recognition
<https://papers.nips.cc/paper/8377-multi-label-co-regularization-for-semi-supervised-facial-action-unit-recognition>
AUTHORS: Xuesong Niu, Hu Han, Shiguang Shan, Xilin Chen
HIGHLIGHT: In this work, we propose a semi-supervised approach for AU recognition utilizing a large number of web face images without AU labels and a small face dataset with AU labels inspired by the co-training methods.

83, **TITLE:** Correlated Uncertainty for Learning Dense Correspondences from Noisy Labels
<https://papers.nips.cc/paper/8378-correlated-uncertainty-for-learning-dense-correspondences-from-noisy-labels>
AUTHORS: Natalia Neverova, David Novotny, Andrea Vedaldi
HIGHLIGHT: We address this issue by augmenting neural network predictors with the ability to output a distribution over labels, thus explicitly and introspectively capturing the aleatoric uncertainty in the annotations.

84, **TITLE:** Powerset Convolutional Neural Networks
<https://papers.nips.cc/paper/8379-powerset-convolutional-neural-networks>
AUTHORS: Chris Wandler, Markus P?schel, Dan Alistarh
HIGHLIGHT: We present a novel class of convolutional neural networks (CNNs) for set functions, i.e., data indexed with the powerset of a finite set.

85, **TITLE:** Optimal Pricing in Repeated Posted-Price Auctions with Different Patience of the Seller and the Buyer
<https://papers.nips.cc/paper/8380-optimal-pricing-in-repeated-posted-price-auctions-with-different-patience-of-the-seller-and-the-buyer>
AUTHORS: Arsenii Vanunts, Alexey Drutsa
HIGHLIGHT: We study revenue optimization pricing algorithms for repeated posted-price auctions where a seller interacts with a single strategic buyer that holds a fixed private valuation.

86, **TITLE:** An Accelerated Decentralized Stochastic Proximal Algorithm for Finite Sums
<https://papers.nips.cc/paper/8381-an-accelerated-decentralized-stochastic-proximal-algorithm-for-finite-sums>
AUTHORS: Hadrien Hendrikx, Francis Bach, Laurent Massouli?
HIGHLIGHT: In this work, we propose an efficient A ccelerated D ecentralized stochastic algorithm for F inite S ums named ADFS, which uses local stochastic proximal updates and randomized pairwise communications between nodes.

87, **TITLE:** Point-Voxel CNN for Efficient 3D Deep Learning
<https://papers.nips.cc/paper/8382-point-voxel-cnn-for-efficient-3d-deep-learning>
AUTHORS: Zhijian Liu, Haotian Tang, Yujun Lin, Song Han
HIGHLIGHT: In this paper, we propose PVCNN that represents the 3D input data in points to reduce the memory consumption, while performing the convolutions in voxels to reduce the irregular, sparse data access and improve the locality.

88, **TITLE:** Deep Learning without Weight Transport
<https://papers.nips.cc/paper/8383-deep-learning-without-weight-transport>
AUTHORS: Mohamed Akrouf, Collin Wilson, Peter Humphreys, Timothy Lillicrap, Douglas B. Tweed
HIGHLIGHT: Here we describe two mechanisms — a neural circuit called a weight mirror and a modification of an algorithm proposed by Kolen and Pollack in 1994 — both of which let the feedback path learn appropriate synaptic weights quickly and accurately even in large networks, without weight transport or complex wiring.

89, **TITLE:** Combinatorial Bandits with Relative Feedback
<https://papers.nips.cc/paper/8384-combinatorial-bandits-with-relative-feedback>
AUTHORS: Aadirupa Saha, Aditya Gopalan
HIGHLIGHT: For both settings, we devise instance-dependent and order-optimal regret algorithms with regret $\mathcal{O}(\frac{n}{m} \ln T)$ and $\mathcal{O}(\frac{n}{k} \ln T)$, respectively.

- 90, TITLE: General Proximal Incremental Aggregated Gradient Algorithms: Better and Novel Results under General Scheme
<https://papers.nips.cc/paper/8385-general-proximal-incremental-aggregated-gradient-algorithms-better-and-novel-results-under-general-scheme>
AUTHORS: Tao Sun, Yuejiao Sun, Dongsheng Li, Qing Liao
HIGHLIGHT: In this paper, we propose a general proximal incremental aggregated gradient algorithm, which contains various existing algorithms including the basic incremental aggregated gradient method.
- 91, TITLE: A Condition Number for Joint Optimization of Cycle-Consistent Networks
<https://papers.nips.cc/paper/8386-a-condition-number-for-joint-optimization-of-cycle-consistent-networks>
AUTHORS: Leonidas J. Guibas, Qixing Huang, Zhenxiao Liang
HIGHLIGHT: This paper presents an algorithm that select a subset of weighted cycles to minimize a condition number of the induced joint optimization problem.
- 92, TITLE: Explicit Disentanglement of Appearance and Perspective in Generative Models
<https://papers.nips.cc/paper/8387-explicit-disentanglement-of-appearance-and-perspective-in-generative-models>
AUTHORS: Nicki Skafte, Søren Hauberg
HIGHLIGHT: Specifically, we propose a model with two latent spaces: one that represents spatial transformations of the input data, and another that represents the transformed data.
- 93, TITLE: Polynomial Cost of Adaptation for X-Armed Bandits
<https://papers.nips.cc/paper/8388-polynomial-cost-of-adaptation-for-x-armed-bandits>
AUTHORS: Hedi Hadji
HIGHLIGHT: In the context of stochastic continuum-armed bandits, we present an algorithm that adapts to the unknown smoothness of the objective function.
- 94, TITLE: Learning to Propagate for Graph Meta-Learning
<https://papers.nips.cc/paper/8389-learning-to-propagate-for-graph-meta-learning>
AUTHORS: LU LIU, Tianyi Zhou, Guodong Long, Jing Jiang, Chengqi Zhang
HIGHLIGHT: In this paper, we show that a meta-learner that explicitly relates tasks on a graph describing the relations of their output dimensions (e.g., classes) can significantly improve few-shot learning.
- 95, TITLE: Secretary Ranking with Minimal Inversions
<https://papers.nips.cc/paper/8390-secretary-ranking-with-minimal-inversions>
AUTHORS: Sepehr Assadi, Eric Balkanski, Renato Leme
HIGHLIGHT: We present an algorithm that ranks n elements with only $O(n^{\{3/2\}})$ inversions in expectation, and show that any algorithm necessarily suffers $\Omega(n^{\{3/2\}})$ inversions when there are n available positions.
- 96, TITLE: Nonparametric Regressive Point Processes Based on Conditional Gaussian Processes
<https://papers.nips.cc/paper/8391-nonparametric-regressive-point-processes-based-on-conditional-gaussian-processes>
AUTHORS: Siqi Liu, Milos Hauskrecht
HIGHLIGHT: In this work, we propose and develop a new nonparametric regressive point process model based on Gaussian processes.
- 97, TITLE: Learning Perceptual Inference by Contrasting
<https://papers.nips.cc/paper/8392-learning-perceptual-inference-by-contrasting>
AUTHORS: Chi Zhang, Baoxiong Jia, Feng Gao, Yixin Zhu, HongJing Lu, Song-Chun Zhu
HIGHLIGHT: In this work, we study how to improve machines' reasoning ability on one challenging task of this kind: Raven's Progressive Matrices (RPM).
- 98, TITLE: Selecting the independent coordinates of manifolds with large aspect ratios
<https://papers.nips.cc/paper/8393-selecting-the-independent-coordinates-of-manifolds-with-large-aspect-ratios>
AUTHORS: Yu-Chia Chen, Marina Meila
HIGHLIGHT: Hence, we propose a bicriterial Independent Eigencoordinate Selection (IES) algorithm that selects smooth embeddings with few eigenvectors.
- 99, TITLE: Region-specific Diffeomorphic Metric Mapping
<https://papers.nips.cc/paper/8394-region-specific-diffeomorphic-metric-mapping>
AUTHORS: Zhengyang Shen, Francois-Xavier Vialard, Marc Niethammer
HIGHLIGHT: We introduce a region-specific diffeomorphic metric mapping (RDMM) registration approach.

100, TITLE: Deep Supervised Summarization: Algorithm and Application to Learning Instructions
<https://papers.nips.cc/paper/8395-deep-supervised-summarization-algorithm-and-application-to-learning-instructions>
AUTHORS: Chengguang Xu, Ehsan Elhamifar
HIGHLIGHT: To do so, we propose to learn representations of data so that the input of transformed data to the facility location recovers their ground-truth representatives.

101, TITLE: Scene Representation Networks: Continuous 3D-Structure-Aware Neural Scene Representations
<https://papers.nips.cc/paper/8396-scene-representation-networks-continuous-3d-structure-aware-neural-scene-representations>
AUTHORS: Vincent Sitzmann, Michael Zollhoefer, Gordon Wetzstein
HIGHLIGHT: We propose Scene Representation Networks (SRNs), a continuous, 3D-structure-aware scene representation that encodes both geometry and appearance.

102, TITLE: Reconciling γ -Returns with Experience Replay
<https://papers.nips.cc/paper/8397-reconciling-returns-with-experience-replay>
AUTHORS: Brett Daley, Christopher Amato
HIGHLIGHT: Towards this, we propose the first method to enable practical use of γ -returns in arbitrary replay-based methods without relying on other forms of decorrelation such as asynchronous gradient updates.

103, TITLE: Control Batch Size and Learning Rate to Generalize Well: Theoretical and Empirical Evidence
<https://papers.nips.cc/paper/8398-control-batch-size-and-learning-rate-to-generalize-well-theoretical-and-empirical-evidence>
AUTHORS: Fengxiang He, Tongliang Liu, Dacheng Tao
HIGHLIGHT: This paper reports both theoretical and empirical evidence of a training strategy that we should control the ratio of batch size to learning rate not too large to achieve a good generalization ability.

104, TITLE: Non-Asymptotic Gap-Dependent Regret Bounds for Tabular MDPs
<https://papers.nips.cc/paper/8399-non-asymptotic-gap-dependent-regret-bounds-for-tabular-mdps>
AUTHORS: Max Simchowitz, Kevin G. Jamieson
HIGHLIGHT: This paper establishes that optimistic algorithms attain gap-dependent and non-asymptotic logarithmic regret for episodic MDPs.

105, TITLE: A Graph Theoretic Framework of Recomputation Algorithms for Memory-Efficient Backpropagation
<https://papers.nips.cc/paper/8400-a-graph-theoretic-framework-of-recomputation-algorithms-for-memory-efficient-backpropagation>
AUTHORS: Mitsuru Kusumoto, Takuya Inoue, Gentaro Watanabe, Takuya Akiba, Masanori Koyama
HIGHLIGHT: In this paper, we will propose a novel and efficient recomputation method that can be applied to a wider range of neural nets than previous methods.

106, TITLE: Combinatorial Inference against Label Noise
<https://papers.nips.cc/paper/8401-combinatorial-inference-against-label-noise>
AUTHORS: Paul Hongsuck Seo, Geeho Kim, Bohyung Han
HIGHLIGHT: To handle the label noise issue in a principled way, we propose a unique classification framework of constructing multiple models in heterogeneous coarse-grained meta-class spaces and making joint inference of the trained models for the final predictions in the original (base) class space.

107, TITLE: Value Propagation for Decentralized Networked Deep Multi-agent Reinforcement Learning
<https://papers.nips.cc/paper/8402-value-propagation-for-decentralized-networked-deep-multi-agent-reinforcement-learning>
AUTHORS: Chao Qu, Shie Mannor, Huan Xu, Yuan Qi, Le Song, Junwu Xiong
HIGHLIGHT: We consider the networked multi-agent reinforcement learning (MARL) problem in a fully decentralized setting, where agents learn to coordinate to achieve joint success.

108, TITLE: Convolution with even-sized kernels and symmetric padding
<https://papers.nips.cc/paper/8403-convolution-with-even-sized-kernels-and-symmetric-padding>
AUTHORS: Shuang Wu, Guanrui Wang, Pei Tang, Feng Chen, Luping Shi
HIGHLIGHT: In this work, we quantify the shift problem occurs in even-sized kernel convolutions by an information erosion hypothesis, and eliminate it by proposing symmetric padding on four sides of the feature maps (C2sp, C4sp).

109, TITLE: On The Classification-Distortion-Perception Tradeoff
<https://papers.nips.cc/paper/8404-on-the-classification-distortion-perception-tradeoff>
AUTHORS: Dong Liu, Haochen Zhang, Zhiwei Xiong

HIGHLIGHT: In this paper, we extend the previous perception-distortion tradeoff to the case of classification-distortion-perception (CDP) tradeoff, where we introduced the classification error rate of the restored signal in addition to distortion and perceptual difference.

110, **TITLE:** Optimal Statistical Rates for Decentralised Non-Parametric Regression with Linear Speed-Up
<https://papers.nips.cc/paper/8405-optimal-statistical-rates-for-decentralised-non-parametric-regression-with-linear-speed-up>
AUTHORS: Dominic Richards, Patrick Rebeschini
HIGHLIGHT: We analyse the learning performance of Distributed Gradient Descent in the context of multi-agent decentralised non-parametric regression with the square loss function when i.i.d. samples are assigned to agents.

111, **TITLE:** Online sampling from log-concave distributions
<https://papers.nips.cc/paper/8406-online-sampling-from-log-concave-distributions>
AUTHORS: Holden Lee, Oren Mangoubi, Nisheeth Vishnoi
HIGHLIGHT: Technically, lack of strong convexity is a significant barrier to analysis and, here, our main contribution is a martingale exit time argument that shows our Markov chain remains in a ball of radius roughly poly-logarithmic in ϵ for enough time to reach within ϵ of π_t .

112, **TITLE:** Envy-Free Classification
<https://papers.nips.cc/paper/8407-envy-free-classification>
AUTHORS: Maria-Florina F. Balcan, Travis Dick, Ritesh Noothigattu, Ariel D. Procaccia
HIGHLIGHT: On a conceptual level, we argue that envy-freeness also provides a compelling notion of fairness for classification tasks, especially when individuals have heterogeneous preferences.

113, **TITLE:** Finding Friend and Foe in Multi-Agent Games
<https://papers.nips.cc/paper/8408-finding-friend-and-foe-in-multi-agent-games>
AUTHORS: Jack Serrino, Max Kleiman-Weiner, David C. Parkes, Josh Tenenbaum
HIGHLIGHT: Here we develop the DeepRole algorithm, a multi-agent reinforcement learning agent that we test on "The Resistance: Avalon", the most popular hidden role game.

114, **TITLE:** Image Synthesis with a Single (Robust) Classifier
<https://papers.nips.cc/paper/8409-image-synthesis-with-a-single-robust-classifier>
AUTHORS: Shibani Santurkar, Andrew Ilyas, Dimitris Tsipras, Logan Engstrom, Brandon Tran, Aleksander Madry
HIGHLIGHT: We show that the basic classification framework alone can be used to tackle some of the most challenging tasks in image synthesis.

115, **TITLE:** Model Compression with Adversarial Robustness: A Unified Optimization Framework
<https://papers.nips.cc/paper/8410-model-compression-with-adversarial-robustness-a-unified-optimization-framework>
AUTHORS: Shupeng Gui, Haotao N. Wang, Haichuan Yang, Chen Yu, Zhangyang Wang, Ji Liu
HIGHLIGHT: We propose a novel Adversarially Trained Model Compression (ATMC) framework.

116, **TITLE:** Cross-channel Communication Networks
<https://papers.nips.cc/paper/8411-cross-channel-communication-networks>
AUTHORS: Jianwei Yang, Zhile Ren, Chuang Gan, Hongyuan Zhu, Devi Parikh
HIGHLIGHT: We introduce a novel network unit called Cross-channel Communication (C3) block, a simple yet effective module to encourage the neuron communication within the same layer.

117, **TITLE:** CondConv: Conditionally Parameterized Convolutions for Efficient Inference
<https://papers.nips.cc/paper/8412-condconv-conditionally-parameterized-convolutions-for-efficient-inference>
AUTHORS: Brandon Yang, Gabriel Bender, Quoc V. Le, Jiquan Ngiam
HIGHLIGHT: We propose conditionally parameterized convolutions (CondConv), which learn specialized convolutional kernels for each example.

118, **TITLE:** Regression Planning Networks
<https://papers.nips.cc/paper/8413-regression-planning-networks>
AUTHORS: Danfei Xu, Roberto Marten, De-An Huang, Yuke Zhu, Silvio Savarese, Li F. Fei-Fei
HIGHLIGHT: In this work, we combine the benefits of these two paradigms and propose a learning-to-plan method that can directly generate a long-term symbolic plan conditioned on high-dimensional observations.

119, **TITLE:** Twin Auxiliary Classifiers GAN
<https://papers.nips.cc/paper/8414-twin-auxiliary-classifiers-gan>

- AUTHORS: Mingming Gong, Yanwu Xu, Chunyuan Li, Kun Zhang, Kayhan Batmanghelich
HIGHLIGHT: In this paper, we identify the source of low diversity issue theoretically and propose a practical solution to the problem.
- 120, TITLE: Conditional Structure Generation through Graph Variational Generative Adversarial Nets
<https://papers.nips.cc/paper/8415-conditional-structure-generation-through-graph-variational-generative-adversarial-nets>
AUTHORS: Carl Yang, Peiye Zhuang, Wenhan Shi, Alan Luu, Pan Li
HIGHLIGHT: While existing graph generative models only consider graph structures without semantic contexts, we formulate the novel problem of conditional structure generation, and propose a novel unified model of graph variational generative adversarial nets (CondGen) to handle the intrinsic challenges of flexible context-structure conditioning and permutation-invariant generation.
- 121, TITLE: Distributional Policy Optimization: An Alternative Approach for Continuous Control
<https://papers.nips.cc/paper/8416-distributional-policy-optimization-an-alternative-approach-for-continuous-control>
AUTHORS: Chen Tessler, Guy Tennenholtz, Shie Mannor
HIGHLIGHT: We identify a fundamental problem in policy gradient-based methods in continuous control.
- 122, TITLE: Sampling Sketches for Concave Sublinear Functions of Frequencies
<https://papers.nips.cc/paper/8417-sampling-sketches-for-concave-sublinear-functions-of-frequencies>
AUTHORS: Edith Cohen, Ofir Geri
HIGHLIGHT: Our main contribution is the design of composable sampling sketches that can be tailored to any concave sublinear function of the frequencies.
- 123, TITLE: Deliberative Explanations: visualizing network insecurities
<https://papers.nips.cc/paper/8418-deliberative-explanations-visualizing-network-insecurities>
AUTHORS: Pei Wang, Nuno Vasconcelos
HIGHLIGHT: A new approach to explainable AI, denoted $\{\text{it deliberative explanations,}\}$ is proposed.
- 124, TITLE: Computing Full Conformal Prediction Set with Approximate Homotopy
<https://papers.nips.cc/paper/8419-computing-full-conformal-prediction-set-with-approximate-homotopy>
AUTHORS: Eugene Ndiaye, Ichiro Takeuchi
HIGHLIGHT: We propose efficient algorithms to compute conformal prediction set using approximated solution of (convex) regularized empirical risk minimization.
- 125, TITLE: Failing Loudly: An Empirical Study of Methods for Detecting Dataset Shift
<https://papers.nips.cc/paper/8420-failing-loudly-an-empirical-study-of-methods-for-detecting-dataset-shift>
AUTHORS: Stephan Rabanser, Stephan G?nnemann, Zachary Lipton
HIGHLIGHT: This paper explores the problem of building ML systems that fail loudly, investigating methods for detecting dataset shift, identifying exemplars that most typify the shift, and quantifying shift malignancy.
- 126, TITLE: Hierarchical Reinforcement Learning with Advantage-Based Auxiliary Rewards
<https://papers.nips.cc/paper/8421-hierarchical-reinforcement-learning-with-advantage-based-auxiliary-rewards>
AUTHORS: Siyuan Li, Rui Wang, Minxue Tang, Chongjie Zhang
HIGHLIGHT: In this paper, we aim to adapt low-level skills to downstream tasks while maintaining the generality of reward design.
- 127, TITLE: Multi-View Reinforcement Learning
<https://papers.nips.cc/paper/8422-multi-view-reinforcement-learning>
AUTHORS: Minne Li, Lisheng Wu, Jun WANG, Haitham Bou Ammar
HIGHLIGHT: We define the MVRL framework by extending partially observable Markov decision processes (POMDPs) to support more than one observation model and propose two solution methods through observation augmentation and cross-view policy transfer.
- 128, TITLE: Cascade RPN: Delving into High-Quality Region Proposal Network with Adaptive Convolution
<https://papers.nips.cc/paper/8423-cascade-rpn-delving-into-high-quality-region-proposal-network-with-adaptive-convolution>
AUTHORS: Thang Vu, Hyunjun Jang, Trung X. Pham, Chang Yoo
HIGHLIGHT: This paper considers an architecture referred to as Cascade Region Proposal Network (Cascade RPN) for improving the region-proposal quality and detection performance by systematically addressing the limitation of the conventional RPN that heuristically defines the anchors and aligns the features to the anchors.
- 129, TITLE: Neural Diffusion Distance for Image Segmentation

<https://papers.nips.cc/paper/8424-neural-diffusion-distance-for-image-segmentation>

AUTHORS: Jian Sun, Zongben Xu

HIGHLIGHT: In this work, we propose a spec-diff-net for computing diffusion distance on graph based on approximate spectral decomposition.

130, TITLE: Fine-grained Optimization of Deep Neural Networks

<https://papers.nips.cc/paper/8425-fine-grained-optimization-of-deep-neural-networks>

AUTHORS: Mete Ozay

HIGHLIGHT: In this work, we conjecture that if we can impose multiple constraints on weights of DNNs to upper bound the norms of the weights, and train the DNNs with these weights, then we can attain empirical generalization errors closer to the derived theoretical bounds, and improve accuracy of the DNNs.

To this end, we pose two problems.

131, TITLE: Extending Stein's unbiased risk estimator to train deep denoisers with correlated pairs of noisy images

<https://papers.nips.cc/paper/8426-extending-steins-unbiased-risk-estimator-to-train-deep-denoisers-with-correlated-pairs-of-noisy-images>

AUTHORS: Magauyi Zhussip, Shakarim Soltanayev, Se Young Chun

HIGHLIGHT: Here, we propose an extended SURE (eSURE) to train deep denoisers with correlated pairs of noise realizations per image and applied it to the case with two uncorrelated realizations per image to achieve better performance than SURE based method and comparable results to Noise2Noise.

132, TITLE: Fixing Implicit Derivatives: Trust-Region Based Learning of Continuous Energy Functions

<https://papers.nips.cc/paper/8427-fixing-implicit-derivatives-trust-region-based-learning-of-continuous-energy-functions>

AUTHORS: Chris Russell, Matteo Toso, Neill Campbell

HIGHLIGHT: We present a new technique for the learning of continuous energy functions that we refer to as Wibergian Learning.

133, TITLE: Hyperspherical Prototype Networks

<https://papers.nips.cc/paper/8428-hyperspherical-prototype-networks>

AUTHORS: Pascal Mettes, Elise van der Pol, Cees Snoek

HIGHLIGHT: This paper introduces hyperspherical prototype networks, which unify classification and regression with prototypes on hyperspherical output spaces.

134, TITLE: Expressive power of tensor-network factorizations for probabilistic modeling

<https://papers.nips.cc/paper/8429-expressive-power-of-tensor-network-factorizations-for-probabilistic-modeling>

AUTHORS: Ivan Glasser, Ryan Sweke, Nicola Pancotti, Jens Eisert, Ignacio Cirac

HIGHLIGHT: Inspired by these developments, and the natural correspondence between tensor networks and probabilistic graphical models, we provide a rigorous analysis of the expressive power of various tensor-network factorizations of discrete multivariate probability distributions.

135, TITLE: HyperGCN: A New Method For Training Graph Convolutional Networks on Hypergraphs

<https://papers.nips.cc/paper/8430-hypergcn-a-new-method-for-training-graph-convolutional-networks-on-hypergraphs>

AUTHORS: Naganand Yadati, Madhav Nimishakavi, Prateek Yadav, Vikram Nitin, Anand Louis, Partha Talukdar

HIGHLIGHT: Motivated by the fact that a graph convolutional network (GCN) has been effective for graph-based SSL, we propose HyperGCN, a novel GCN for SSL on attributed hypergraphs.

136, TITLE: SSRGD: Simple Stochastic Recursive Gradient Descent for Escaping Saddle Points

<https://papers.nips.cc/paper/8431-ssrgd-simple-stochastic-recursive-gradient-descent-for-escaping-saddle-points>

AUTHORS: Zhize Li

HIGHLIGHT: We analyze stochastic gradient algorithms for optimizing nonconvex problems.

137, TITLE: Efficient Meta Learning via Minibatch Proximal Update

<https://papers.nips.cc/paper/8432-efficient-meta-learning-via-minibatch-proximal-update>

AUTHORS: Pan Zhou, Xiaotong Yuan, Huan Xu, Shuicheng Yan, Jiashi Feng

HIGHLIGHT: To remedy this deficiency, in this paper we propose a minibatch proximal update based meta-learning approach for learning to efficient hypothesis transfer.

138, TITLE: Unconstrained Monotonic Neural Networks

<https://papers.nips.cc/paper/8433-unconstrained-monotonic-neural-networks>

AUTHORS: Antoine Wehenkel, Gilles Louppe

- HIGHLIGHT:** In this work, we propose the Unconstrained Monotonic Neural Network (UMNN) architecture based on the insight that a function is monotonic as long as its derivative is strictly positive.
- 139, **TITLE:** Guided Similarity Separation for Image Retrieval
<https://papers.nips.cc/paper/8434-guided-similarity-separation-for-image-retrieval>
AUTHORS: Chundi Liu, Guangwei Yu, Maksims Volkovs, Cheng Chang, Himanshu Rai, Junwei Ma, Satya Krishna Gorti
HIGHLIGHT: In this work we propose a different approach where we leverage graph convolutional networks to directly encode neighbor information into image descriptors.
- 140, **TITLE:** Learning Imbalanced Datasets with Label-Distribution-Aware Margin Loss
<https://papers.nips.cc/paper/8435-learning-imbalanced-datasets-with-label-distribution-aware-margin-loss>
AUTHORS: Kaidi Cao, Colin Wei, Adrien Gaidon, Nikos Arechiga, Tengyu Ma
HIGHLIGHT: Deep learning algorithms can fare poorly when the training dataset suffers from heavy class-imbalance but the testing criterion requires good generalization on less frequent classes. We design two novel methods to improve performance in such scenarios.
- 141, **TITLE:** Strategizing against No-regret Learners
<https://papers.nips.cc/paper/8436-strategizing-against-no-regret-learners>
AUTHORS: Yuan Deng, Jon Schneider, Balasubramanian Sivan
HIGHLIGHT: We study this question and show that under some mild assumptions, the player can always guarantee himself a utility of at least what he would get in a Stackelberg equilibrium.
- 142, **TITLE:** D-VAE: A Variational Autoencoder for Directed Acyclic Graphs
<https://papers.nips.cc/paper/8437-d-vae-a-variational-autoencoder-for-directed-acyclic-graphs>
AUTHORS: Muhan Zhang, Shali Jiang, Zhicheng Cui, Roman Garnett, Yixin Chen
HIGHLIGHT: In this paper, we study deep generative models for DAGs, and propose a novel DAG variational autoencoder (D-VAE).
- 143, **TITLE:** Hierarchical Optimal Transport for Document Representation
<https://papers.nips.cc/paper/8438-hierarchical-optimal-transport-for-document-representation>
AUTHORS: Mikhail Yurochkin, Sebastian Claiçi, Edward Chien, Farzaneh Mirzazadeh, Justin M. Solomon
HIGHLIGHT: As an alternative, we introduce hierarchical optimal transport as a meta-distance between documents, where documents are modeled as distributions over topics, which themselves are modeled as distributions over words.
- 144, **TITLE:** Multivariate Sparse Coding of Nonstationary Covariances with Gaussian Processes
<https://papers.nips.cc/paper/8439-multivariate-sparse-coding-of-nonstationary-covariances-with-gaussian-processes>
AUTHORS: Rui Li
HIGHLIGHT: We propose a unified nonstationary modeling framework to jointly encode the observation correlations to generate a piece-wise representation with a hyper-level Gaussian process (GP) governing the overall contour of the pieces.
- 145, **TITLE:** Positional Normalization
<https://papers.nips.cc/paper/8440-positional-normalization>
AUTHORS: Boyi Li, Felix Wu, Kilian Q. Weinberger, Serge Belongie
HIGHLIGHT: In this paper, we propose a novel normalization method that deviates from this theme.
- 146, **TITLE:** A New Defense Against Adversarial Images: Turning a Weakness into a Strength
<https://papers.nips.cc/paper/8441-a-new-defense-against-adversarial-images-turning-a-weakness-into-a-strength>
AUTHORS: Shengyuan Hu, Tao Yu, Chuan Guo, Wei-Lun Chao, Kilian Q. Weinberger
HIGHLIGHT: In this paper, we adopt a novel perspective and regard the omnipresence of adversarial perturbations as a strength rather than a weakness.
- 147, **TITLE:** Quadratic Video Interpolation
<https://papers.nips.cc/paper/8442-quadratic-video-interpolation>
AUTHORS: Xiangyu Xu, Li Siyao, Wenxiu Sun, Qian Yin, Ming-Hsuan Yang
HIGHLIGHT: To address these issues, we propose a quadratic video interpolation method which exploits the acceleration information in videos.
- 148, **TITLE:** ResNets Ensemble via the Feynman-Kac Formalism to Improve Natural and Robust Accuracies
<https://papers.nips.cc/paper/8443-resnets-ensemble-via-the-feynman-kac-formalism-to-improve-natural-and-robust-accuracies>
AUTHORS: Bao Wang, Zuoqiang Shi, Stanley Osher

HIGHLIGHT: Based on this unified viewpoint, we propose a simple yet effective ResNets ensemble algorithm to boost the accuracy of the robustly trained model on both clean and adversarial images.

149, **TITLE:** Incremental Scene Synthesis
<https://papers.nips.cc/paper/8444-incremental-scene-synthesis>
AUTHORS: Benjamin Planche, Xuejian Rong, Ziyang Wu, Srikrishna Karanam, Harald Kosch, YingLi Tian, Jan Ernst, ANDREAS HUTTER
HIGHLIGHT: We present a method to incrementally generate complete 2D or 3D scenes with the following properties: (a) it is globally consistent at each step according to a learned scene prior, (b) real observations of a scene can be incorporated while observing global consistency, (c) unobserved regions can be hallucinated locally in consistence with previous observations, hallucinations and global priors, and (d) hallucinations are statistical in nature, i.e., different scenes can be generated from the same observations.

150, **TITLE:** Self-Supervised Generalisation with Meta Auxiliary Learning
<https://papers.nips.cc/paper/8445-self-supervised-generalisation-with-meta-auxiliary-learning>
AUTHORS: Shikun Liu, Andrew Davison, Edward Johns
HIGHLIGHT: We propose a new method which automatically learns appropriate labels for an auxiliary task, such that any supervised learning task can be improved without requiring access to any further data.

151, **TITLE:** Variational Denoising Network: Toward Blind Noise Modeling and Removal
<https://papers.nips.cc/paper/8446-variational-denoising-network-toward-blind-noise-modeling-and-removal>
AUTHORS: Zongsheng Yue, Hongwei Yong, Qian Zhao, Deyu Meng, Lei Zhang
HIGHLIGHT: In this work we propose a new variational inference method, which integrates both noise estimation and image denoising into a unique Bayesian framework, for blind image denoising.

152, **TITLE:** Fast Sparse Group Lasso
<https://papers.nips.cc/paper/8447-fast-sparse-group-lasso>
AUTHORS: Yasutoshi Ida, Yasuhiro Fujiwara, Hisashi Kashima
HIGHLIGHT: This paper proposes a fast Block Coordinate Descent for Sparse Group Lasso.

153, **TITLE:** Learnable Tree Filter for Structure-preserving Feature Transform
<https://papers.nips.cc/paper/8448-learnable-tree-filter-for-structure-preserving-feature-transform>
AUTHORS: Lin Song, Yanwei Li, Zeming Li, Gang Yu, Hongbin Sun, Jian Sun, Nanning Zheng
HIGHLIGHT: In this paper, we propose the learnable tree filter to form a generic tree filtering module that leverages the structural property of minimal spanning tree to model long-range dependencies while preserving the details.

154, **TITLE:** Data-Dependence of Plateau Phenomenon in Learning with Neural Network --- Statistical Mechanical Analysis
<https://papers.nips.cc/paper/8449-data-dependence-of-plateau-phenomenon-in-learning-with-neural-network-statistical-mechanical-analysis>
AUTHORS: Yuki Yoshida, Masato Okada
HIGHLIGHT: In this paper, using statistical mechanical formulation, we clarified the relationship between the plateau phenomenon and the statistical property of the data learned.

155, **TITLE:** Coordinated hippocampal-entorhinal replay as structural inference
<https://papers.nips.cc/paper/8450-coordinated-hippocampal-entorhinal-replay-as-structural-inference>
AUTHORS: Talfan Evans, Neil Burgess
HIGHLIGHT: We propose that this offline inference corresponds to coordinated hippocampal-entorhinal replay during sharp wave ripples.

156, **TITLE:** Cascaded Dilated Dense Network with Two-step Data Consistency for MRI Reconstruction
<https://papers.nips.cc/paper/8451-cascaded-dilated-dense-network-with-two-step-data-consistency-for-mri-reconstruction>
AUTHORS: Hao Zheng, Faming Fang, Guixu Zhang
HIGHLIGHT: Inspired by recent deep learning methods, we propose a Cascaded Dilated Dense Network (CDDN) for MRI reconstruction.

157, **TITLE:** On the Ineffectiveness of Variance Reduced Optimization for Deep Learning
<https://papers.nips.cc/paper/8452-on-the-ineffectiveness-of-variance-reduced-optimization-for-deep-learning>
AUTHORS: Aaron Defazio, Leon Bottou
HIGHLIGHT: We show that naive application of the SVRG technique and related approaches fail, and explore why.

158, **TITLE:** On the Curved Geometry of Accelerated Optimization

<https://papers.nips.cc/paper/8453-on-the-curved-geometry-of-accelerated-optimization>
AUTHORS: Aaron Defazio
HIGHLIGHT: In this work we propose a differential geometric motivation for Nesterov's accelerated gradient method (AGM) for strongly-convex problems.

159, TITLE: Multi-marginal Wasserstein GAN
<https://papers.nips.cc/paper/8454-multi-marginal-wasserstein-gan>
AUTHORS: Jiezhong Cao, Langyuan Mo, Yifan Zhang, Kui Jia, Chunhua Shen, Mingkui Tan
HIGHLIGHT: In this paper, we propose a novel Multi-marginal Wasserstein GAN (MWGAN) to minimize Wasserstein distance among domains.

160, TITLE: Better Exploration with Optimistic Actor Critic
<https://papers.nips.cc/paper/8455-better-exploration-with-optimistic-actor-critic>
AUTHORS: Kamil Ciosek, Quan Vuong, Robert Loftin, Katja Hofmann
HIGHLIGHT: To address both of these phenomena, we introduce a new algorithm, Optimistic Actor Critic, which approximates a lower and upper confidence bound on the state-action value function.

161, TITLE: Importance Resampling for Off-policy Prediction
<https://papers.nips.cc/paper/8456-importance-resampling-for-off-policy-prediction>
AUTHORS: Matthew Schlegel, Wesley Chung, Daniel Graves, Jian Qian, Martha White
HIGHLIGHT: In this work, we explore a resampling strategy as an alternative to reweighting.

162, TITLE: The Label Complexity of Active Learning from Observational Data
<https://papers.nips.cc/paper/8457-the-label-complexity-of-active-learning-from-observational-data>
AUTHORS: Songbai Yan, Kamalika Chaudhuri, Tara Javidi
HIGHLIGHT: We provably demonstrate that the result of this is an algorithm which is statistically consistent as well as more label-efficient than prior work.

163, TITLE: Meta-Learning Representations for Continual Learning
<https://papers.nips.cc/paper/8458-meta-learning-representations-for-continual-learning>
AUTHORS: Khurram Javed, Martha White
HIGHLIGHT: In this paper, we propose OML, an objective that directly minimizes catastrophic interference by learning representations that accelerate future learning and are robust to forgetting under online updates in continual learning.

164, TITLE: Defense Against Adversarial Attacks Using Feature Scattering-based Adversarial Training
<https://papers.nips.cc/paper/8459-defense-against-adversarial-attacks-using-feature-scattering-based-adversarial-training>
AUTHORS: Haichao Zhang, Jianyu Wang
HIGHLIGHT: We introduce a feature scattering-based adversarial training approach for improving model robustness against adversarial attacks.

165, TITLE: Visualizing the PHATE of Neural Networks
<https://papers.nips.cc/paper/8460-visualizing-the-phate-of-neural-networks>
AUTHORS: Scott Gigante, Adam S. Charles, Smita Krishnaswamy, Gal Mishne
HIGHLIGHT: To this end, we introduce a novel visualization algorithm that reveals the internal geometry of such networks: Multislice PHATE (M-PHATE), the first method designed explicitly to visualize how a neural network's hidden representations of data evolve throughout the course of training.

166, TITLE: The Cells Out of Sample (COOS) dataset and benchmarks for measuring out-of-sample generalization of image classifiers
<https://papers.nips.cc/paper/8461-the-cells-out-of-sample-coos-dataset-and-benchmarks-for-measuring-out-of-sample-generalization-of-image-classifiers>
AUTHORS: Alex Lu, Amy Lu, Wiebke Schormann, Marzyeh Ghassemi, David Andrews, Alan Moses
HIGHLIGHT: We created a public dataset of 132,209 images of mouse cells, COOS-7 (Cells Out Of Sample 7-Class).

167, TITLE: Nonconvex Low-Rank Tensor Completion from Noisy Data
<https://papers.nips.cc/paper/8462-nonconvex-low-rank-tensor-completion-from-noisy-data>
AUTHORS: Changxiao Cai, Gen Li, H. Vincent Poor, Yuxin Chen
HIGHLIGHT: Focusing on "incoherent" and well-conditioned tensors of a constant CP rank, we propose a two-stage nonconvex algorithm --- (vanilla) gradient descent following a rough initialization --- that achieves the best of both worlds.

- 168, TITLE: Beyond Online Balanced Descent: An Optimal Algorithm for Smoothed Online Optimization
<https://papers.nips.cc/paper/8463-beyond-online-balanced-descent-an-optimal-algorithm-for-smoothed-online-optimization>
AUTHORS: Gautam Goel, Yiheng Lin, Haoyuan Sun, Adam Wierman
HIGHLIGHT: No existing algorithms have competitive ratios matching this bound, and we show that the state-of-the-art algorithm, Online Balanced Decent (OBD), has a competitive ratio that is $\Omega(m^{-2/3})$.
- 169, TITLE: Channel Gating Neural Networks
<https://papers.nips.cc/paper/8464-channel-gating-neural-networks>
AUTHORS: Weizhe Hua, Yuan Zhou, Christopher M. De Sa, Zhiru Zhang, G. Edward Suh
HIGHLIGHT: This paper introduces channel gating, a dynamic, fine-grained, and hardware-efficient pruning scheme to reduce the computation cost for convolutional neural networks (CNNs).
- 170, TITLE: Neural networks grown and self-organized by noise
<https://papers.nips.cc/paper/8465-neural-networks-grown-and-self-organized-by-noise>
AUTHORS: Guruprasad Raghavan, Matt Thomson
HIGHLIGHT: In this paper, we propose a biologically inspired developmental algorithm that can ‘grow’ a functional, layered neural network from a single initial cell.
- 171, TITLE: Catastrophic Forgetting Meets Negative Transfer: Batch Spectral Shrinkage for Safe Transfer Learning
<https://papers.nips.cc/paper/8466-catastrophic-forgetting-meets-negative-transfer-batch-spectral-shrinkage-for-safe-transfer-learning>
AUTHORS: Xinyang Chen, Sinan Wang, Bo Fu, Mingsheng Long, Jianmin Wang
HIGHLIGHT: In this paper, we launch an in-depth empirical investigation into negative transfer in fine-tuning and find that, for the weight parameters and feature representations, transferability of their spectral components is diverse.
- 172, TITLE: Meta-Weight-Net: Learning an Explicit Mapping For Sample Weighting
<https://papers.nips.cc/paper/8467-meta-weight-net-learning-an-explicit-mapping-for-sample-weighting>
AUTHORS: Jun Shu, Qi Xie, Lixuan Yi, Qian Zhao, Sanping Zhou, Zongben Xu, Deyu Meng
HIGHLIGHT: To address this issue, we propose a method capable of adaptively learning an explicit weighting function directly from data.
- 173, TITLE: Variational Structured Semantic Inference for Diverse Image Captioning
<https://papers.nips.cc/paper/8468-variational-structured-semantic-inference-for-diverse-image-captioning>
AUTHORS: Fuhai Chen, Rongrong Ji, Jiayi Ji, Xiaoshuai Sun, Baochang Zhang, Xuri Ge, Yongjian Wu, Feiyue Huang, Yan Wang
HIGHLIGHT: To model these two inherent diversities in image captioning, we propose a Variational Structured Semantic Inferring model (termed VSSI-cap) executed in a novel structured encoder-inferer-decoder schema.
- 174, TITLE: Mapping State Space using Landmarks for Universal Goal Reaching
<https://papers.nips.cc/paper/8469-mapping-state-space-using-landmarks-for-universal-goal-reaching>
AUTHORS: Zhiao Huang, Hao Su, Fangchen Liu
HIGHLIGHT: We propose a method to address this issue in large MDPs with sparse rewards, in which exploration and routing across remote states are both extremely challenging.
- 175, TITLE: Transferable Normalization: Towards Improving Transferability of Deep Neural Networks
<https://papers.nips.cc/paper/8470-transferable-normalization-towards-improving-transferability-of-deep-neural-networks>
AUTHORS: Ximei Wang, Ying Jin, Mingsheng Long, Jianmin Wang, Michael I. Jordan
HIGHLIGHT: In this paper, we delve into the components of DNN architectures and propose Transferable Normalization (TransNorm) in place of existing normalization techniques.
- 176, TITLE: Random deep neural networks are biased towards simple functions
<https://papers.nips.cc/paper/8471-random-deep-neural-networks-are-biased-towards-simple-functions>
AUTHORS: Giacomo De Palma, Bobak Kiani, Seth Lloyd
HIGHLIGHT: We prove that the binary classifiers of bit strings generated by random wide deep neural networks with ReLU activation function are biased towards simple functions.
- 177, TITLE: XNAS: Neural Architecture Search with Expert Advice
<https://papers.nips.cc/paper/8472-xnas-neural-architecture-search-with-expert-advice>
AUTHORS: Niv Nayman, Asaf Noy, Tal Ridnik, Itamar Friedman, Rong Jin, Lihi Zelnik
HIGHLIGHT: This paper introduces a novel optimization method for differential neural architecture search, based on the theory of prediction with expert advice.

- 178, TITLE: CNN²: Viewpoint Generalization via a Binocular Vision
<https://papers.nips.cc/paper/8473-cnn2-viewpoint-generalization-via-a-binocular-vision>
AUTHORS: Wei-Da Chen, Shan-Hung (Brandon) Wu
HIGHLIGHT: Observing that humans use binocular vision to understand the world, we study in this paper whether the 3D viewpoint generalizability of CNNs can be achieved via a binocular vision.
- 179, TITLE: Generalized Off-Policy Actor-Critic
<https://papers.nips.cc/paper/8474-generalized-off-policy-actor-critic>
AUTHORS: Shangtong Zhang, Wendelin Boehmer, Shimon Whiteson
HIGHLIGHT: We propose a new objective, the counterfactual objective, unifying existing objectives for off-policy policy gradient algorithms in the continuing reinforcement learning (RL) setting.
- 180, TITLE: DAC: The Double Actor-Critic Architecture for Learning Options
<https://papers.nips.cc/paper/8475-dac-the-double-actor-critic-architecture-for-learning-options>
AUTHORS: Shangtong Zhang, Shimon Whiteson
HIGHLIGHT: We apply an actor-critic algorithm on each augmented MDP, yielding the Double Actor-Critic (DAC) architecture.
- 181, TITLE: Numerically Accurate Hyperbolic Embeddings Using Tiling-Based Models
<https://papers.nips.cc/paper/8476-numerically-accurate-hyperbolic-embeddings-using-tiling-based-models>
AUTHORS: Tao Yu, Christopher M. De Sa
HIGHLIGHT: To address this, we propose a new model which uses an integer-based tiling to represent $\{any\}$ point in hyperbolic space with provably bounded numerical error.
- 182, TITLE: Controlling Neural Level Sets
<https://papers.nips.cc/paper/8477-controlling-neural-level-sets>
AUTHORS: Matan Atzmon, Niv Haim, Lior Yariv, Ofer Israelov, Haggai Maron, Yaron Lipman
HIGHLIGHT: In this paper we present a simple and scalable approach to directly control level sets of a deep neural network.
- 183, TITLE: Blended Matching Pursuit
<https://papers.nips.cc/paper/8478-blended-matching-pursuit>
AUTHORS: Cyrille Combettes, Sebastian Pokutta
HIGHLIGHT: We present a blended matching pursuit algorithm, combining coordinate descent-like steps with stronger gradient descent steps, for minimizing a smooth convex function over a linear space spanned by a set of atoms.
- 184, TITLE: An Improved Analysis of Training Over-parameterized Deep Neural Networks
<https://papers.nips.cc/paper/8479-an-improved-analysis-of-training-over-parameterized-deep-neural-networks>
AUTHORS: Difan Zou, Quanquan Gu
HIGHLIGHT: In this paper, we provide an improved analysis of the global convergence of (stochastic) gradient descent for training deep neural networks, which only requires a milder over-parameterization condition than previous work in terms of the training sample size and other problem-dependent parameters.
- 185, TITLE: Controllable Text-to-Image Generation
<https://papers.nips.cc/paper/8480-controllable-text-to-image-generation>
AUTHORS: Bowen Li, Xiaojuan Qi, Thomas Lukasiewicz, Philip Torr
HIGHLIGHT: In this paper, we propose a novel controllable text-to-image generative adversarial network (ControlGAN), which can effectively synthesise high-quality images and also control parts of the image generation according to natural language descriptions.
- 186, TITLE: Improving Textual Network Learning with Variational Homophilic Embeddings
<https://papers.nips.cc/paper/8481-improving-textual-network-learning-with-variational-homophilic-embeddings>
AUTHORS: Wenlin Wang, Chenyang Tao, Zhe Gan, Guoyin Wang, Liqun Chen, Xinyuan Zhang, Ruiyi Zhang, Qian Yang, Ricardo Henao, Lawrence Carin
HIGHLIGHT: Different from most existing methods that optimize a discriminative objective, we introduce Variational Homophilic Embedding (VHE), a fully generative model that learns network embeddings by modeling the semantic (textual) information with a variational autoencoder, while accounting for the structural (topology) information through a novel homophilic prior design.
- 187, TITLE: Rethinking Generative Mode Coverage: A Pointwise Guaranteed Approach
<https://papers.nips.cc/paper/8482-rethinking-generative-mode-coverage-a-pointwise-guaranteed-approach>

AUTHORS: Peilin Zhong, Yuchen Mo, Chang Xiao, Pengyu Chen, Changxi Zheng
HIGHLIGHT: Rethinking this problem from a game-theoretic perspective, we show that a complete mode coverage is firmly attainable.

188, TITLE: The Randomized Midpoint Method for Log-Concave Sampling
<https://papers.nips.cc/paper/8483-the-randomized-midpoint-method-for-log-concave-sampling>
AUTHORS: Ruoqi Shen, Yin Tat Lee
HIGHLIGHT: To solve the sampling problem, we propose a new framework to discretize stochastic differential equations.

189, TITLE: Sample-Efficient Deep Reinforcement Learning via Episodic Backward Update
<https://papers.nips.cc/paper/8484-sample-efficient-deep-reinforcement-learning-via-episodic-backward-update>
AUTHORS: Su Young Lee, Choi Sungik, Sae-Young Chung
HIGHLIGHT: We propose Episodic Backward Update (EBU) - a novel deep reinforcement learning algorithm with a direct value propagation.

190, TITLE: Fully Neural Network based Model for General Temporal Point Processes
<https://papers.nips.cc/paper/8485-fully-neural-network-based-model-for-general-temporal-point-processes>
AUTHORS: Takahiro Omi, naonori ueda, Kazuyuki Aihara
HIGHLIGHT: We herein propose a novel RNN based model in which the time course of the intensity function is represented in a general manner.

191, TITLE: Gate Decorator: Global Filter Pruning Method for Accelerating Deep Convolutional Neural Networks
<https://papers.nips.cc/paper/8486-gate-decorator-global-filter-pruning-method-for-accelerating-deep-convolutional-neural-networks>
AUTHORS: Zhonghui You, Kun Yan, Jinnian Ye, Meng Ma, Ping Wang
HIGHLIGHT: In this work, we propose a global filter pruning algorithm called Gate Decorator, which transforms a vanilla CNN module by multiplying its output by the channel-wise scaling factors (i.e. gate).

192, TITLE: Discrimination in Online Markets: Effects of Social Bias on Learning from Reviews and Policy Design
<https://papers.nips.cc/paper/8487-discrimination-in-online-markets-effects-of-social-bias-on-learning-from-reviews-and-policy-design>
AUTHORS: Faidra Georgia Monachou, Itai Ashlagi
HIGHLIGHT: We study this problem using a two-sided large market model with employers and workers mediated by a platform.

193, TITLE: Provably Powerful Graph Networks
<https://papers.nips.cc/paper/8488-provably-powerful-graph-networks>
AUTHORS: Haggai Maron, Heli Ben-Hamu, Hadar Serviansky, Yaron Lipman
HIGHLIGHT: Differently put, we suggest a simple model that interleaves applications of standard Multilayer-Perceptron (MLP) applied to the feature dimension and matrix multiplication.

194, TITLE: Order Optimal One-Shot Distributed Learning
<https://papers.nips.cc/paper/8489-order-optimal-one-shot-distributed-learning>
AUTHORS: Arsalan Sharifnassab, Saber Salehkaleybar, S. Jamaloddin Golestani
HIGHLIGHT: We propose an algorithm called Multi-Resolution Estimator (MRE) whose expected error is no larger than $\tilde{O}(m^{-1/\max\{d,2\}}n^{-1/2})$, where d is the dimension of the parameter space.

195, TITLE: Information Competing Process for Learning Diversified Representations
<https://papers.nips.cc/paper/8490-information-competing-process-for-learning-diversified-representations>
AUTHORS: Jie Hu, Rongrong Ji, ShengChuan Zhang, Xiaoshuai Sun, Qixiang Ye, Chia-Wen Lin, Qi Tian
HIGHLIGHT: Towards learning diversified representations, a new approach, termed Information Competing Process (ICP), is proposed in this paper.

196, TITLE: GENO -- GENeric Optimization for Classical Machine Learning
<https://papers.nips.cc/paper/8491-geno-generic-optimization-for-classical-machine-learning>
AUTHORS: Soeren Laue, Matthias Mitterreiter, Joachim Giesen
HIGHLIGHT: We show on a wide variety of classical but also some recently suggested problems that the automatically generated solvers are (1) as efficient as well engineered, specialized solvers, (2) more efficient by a decent margin than recent state-of-the-art solvers, and (3) orders of magnitude more efficient than classical modeling language plus solver approaches.

197, TITLE: Conditional Independence Testing using Generative Adversarial Networks
<https://papers.nips.cc/paper/8492-conditional-independence-testing-using-generative-adversarial-networks>

- AUTHORS: Alexis Bellot, Mihaela van der Schaar
HIGHLIGHT: Our contribution is a new test statistic based on samples from a generative adversarial network designed to approximate directly a conditional distribution that encodes the null hypothesis, in a manner that maximizes power (the rate of true negatives).
- 198, TITLE: Online Stochastic Shortest Path with Bandit Feedback and Unknown Transition Function
<https://papers.nips.cc/paper/8493-online-stochastic-shortest-path-with-bandit-feedback-and-unknown-transition-function>
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.
- 199, TITLE: Partitioning Structure Learning for Segmented Linear Regression Trees
<https://papers.nips.cc/paper/8494-partitioning-structure-learning-for-segmented-linear-regression-trees>
AUTHORS: Xiangyu Zheng, Song Xi Chen
HIGHLIGHT: This paper proposes a partitioning structure learning method for segmented linear regression trees (SLRT), which assigns linear predictors over the terminal nodes.
- 200, TITLE: A Tensorized Transformer for Language Modeling
<https://papers.nips.cc/paper/8495-a-tensorized-transformer-for-language-modeling>
AUTHORS: Xindian Ma, Peng Zhang, Shuai Zhang, Nan Duan, Yuexian Hou, Ming Zhou, Dawei Song
HIGHLIGHT: In this paper, based on the ideas of tensor decomposition and parameters sharing, we propose a novel self-attention model (namely Multi-linear attention) with Block-Term Tensor Decomposition (BTD).
- 201, TITLE: Kernel Stein Tests for Multiple Model Comparison
<https://papers.nips.cc/paper/8496-kernel-stein-tests-for-multiple-model-comparison>
AUTHORS: Jen Ning Lim, Makoto Yamada, Bernhard Schölkopf, Wittawat Jitkrittum
HIGHLIGHT: We address the problem of non-parametric multiple model comparison: given S candidate models, decide whether each candidate is as good as the best one(s) or worse than it.
- 202, TITLE: Disentangled behavioural representations
<https://papers.nips.cc/paper/8497-disentangled-behavioural-representations>
AUTHORS: Amir Dezfouli, Hassan Ashtiani, Omar Ghattas, Richard Nock, Peter Dayan, Cheng Soon Ong
HIGHLIGHT: To achieve this, we propose a novel end-to-end learning framework in which an encoder is trained to map the behavior of subjects into a low-dimensional latent space.
- 203, TITLE: More Is Less: Learning Efficient Video Representations by Big-Little Network and Depthwise Temporal Aggregation
<https://papers.nips.cc/paper/8498-more-is-less-learning-efficient-video-representations-by-big-little-network-and-depthwise-temporal-aggregation>
AUTHORS: Quanfu Fan, Chun-Fu (Richard) Chen, Hilde Kuehne, Marco Pistoia, David Cox
HIGHLIGHT: To address this problem, we present an lightweight and memory-friendly architecture for action recognition that performs on par with or better than current architectures by using only a fraction of resources.
- 204, TITLE: Rethinking the CSC Model for Natural Images
<https://papers.nips.cc/paper/8499-rethinking-the-csc-model-for-natural-images>
AUTHORS: Dror Simon, Michael Elad
HIGHLIGHT: In this work we provide new insights regarding the CSC model and its capability to represent natural images, and suggest a Bayesian connection between this model and its patch-based ancestor.
- 205, TITLE: Integrating Bayesian and Discriminative Sparse Kernel Machines for Multi-class Active Learning
<https://papers.nips.cc/paper/8500-integrating-bayesian-and-discriminative-sparse-kernel-machines-for-multi-class-active-learning>
AUTHORS: Weishi Shi, Qi Yu
HIGHLIGHT: We propose a novel active learning (AL) model that integrates Bayesian and discriminative kernel machines for fast and accurate multi-class data sampling.
- 206, TITLE: Learning to Control Self-Assembling Morphologies: A Study of Generalization via Modularity
<https://papers.nips.cc/paper/8501-learning-to-control-self-assembling-morphologies-a-study-of-generalization-via-modularity>
AUTHORS: Deepak Pathak, Christopher Lu, Trevor Darrell, Phillip Isola, Alexei A. Efros
HIGHLIGHT: In contrast, this paper investigates a modular co-evolution strategy: a collection of primitive agents learns to dynamically self-assemble into composite bodies while also learning to coordinate their behavior to control these bodies.

207, TITLE: Perceiving the arrow of time in autoregressive motion
<https://papers.nips.cc/paper/8502-perceiving-the-arrow-of-time-in-autoregressive-motion>
AUTHORS: Kristof Meding, Dominik Janzing, Bernhard Schölkopf, Felix A. Wichmann
HIGHLIGHT: We employ a so-called frozen noise paradigm enabling us to compare human performance with four different algorithms on a trial-by-trial basis: A causal inference algorithm exploiting the dependence structure of additive noise terms, a neurally inspired network, a Bayesian ideal observer model as well as a simple heuristic.

208, TITLE: DualDICE: Behavior-Agnostic Estimation of Discounted Stationary Distribution Corrections
<https://papers.nips.cc/paper/8503-dualdice-behavior-agnostic-estimation-of-discounted-stationary-distribution-corrections>
AUTHORS: Ofir Nachum, Yinlam Chow, Bo Dai, Lihong Li
HIGHLIGHT: In this work, we propose an algorithm, DualDICE, for estimating these quantities.

209, TITLE: Hyper-Graph-Network Decoders for Block Codes
<https://papers.nips.cc/paper/8504-hyper-graph-network-decoders-for-block-codes>
AUTHORS: Eliya Nachmani, Lior Wolf
HIGHLIGHT: In this work, we extend these results to much larger families of algebraic block codes, by performing message passing with graph neural networks.

210, TITLE: Large Scale Markov Decision Processes with Changing Rewards
<https://papers.nips.cc/paper/8505-large-scale-markov-decision-processes-with-changing-rewards>
AUTHORS: Adrian Rivera Cardoso, He Wang, Huan Xu
HIGHLIGHT: By approximating the state-action occupancy measures with a linear architecture of dimension $d|S|$, we propose a modified algorithm with a computational complexity polynomial in d and independent of $|S|$.

211, TITLE: Multiview Aggregation for Learning Category-Specific Shape Reconstruction
<https://papers.nips.cc/paper/8506-multiview-aggregation-for-learning-category-specific-shape-reconstruction>
AUTHORS: Srinath Sridhar, Davis Rempe, Julien Valentin, Bouaziz Sofien, Leonidas J. Guibas
HIGHLIGHT: We present a method that can estimate dense 3D shape, and aggregate shape across multiple and varying number of input views.

212, TITLE: Semi-Parametric Dynamic Contextual Pricing
<https://papers.nips.cc/paper/8507-semi-parametric-dynamic-contextual-pricing>
AUTHORS: Virag Shah, Ramesh Johari, Jose Blanchet
HIGHLIGHT: Motivated by the application of real-time pricing in e-commerce platforms, we consider the problem of revenue-maximization in a setting where the seller can leverage contextual information describing the customer's history and the product's type to predict her valuation of the product.

213, TITLE: Interlaced Greedy Algorithm for Maximization of Submodular Functions in Nearly Linear Time
<https://papers.nips.cc/paper/8508-interlaced-greedy-algorithm-for-maximization-of-submodular-functions-in-nearly-linear-time>
AUTHORS: Alan Kuhnle
HIGHLIGHT: A deterministic approximation algorithm is presented for the maximization of non-monotone submodular functions over a ground set of size n subject to cardinality constraint k ; the algorithm is based upon the idea of interlacing two greedy procedures.

214, TITLE: Initialization of ReLUs for Dynamical Isometry
<https://papers.nips.cc/paper/8509-initialization-of-relus-for-dynamical-isometry>
AUTHORS: Rebekka Burkholz, Alina Dubatovka
HIGHLIGHT: We derive the joint signal output distribution exactly, without mean field assumptions, for fully-connected networks with Gaussian weights and biases, and analyze deviations from the mean field results.

215, TITLE: Gradient Information for Representation and Modeling
<https://papers.nips.cc/paper/8510-gradient-information-for-representation-and-modeling>
AUTHORS: Jie Ding, Robert Calderbank, Vahid Tarokh
HIGHLIGHT: Motivated by Fisher divergence, in this paper we present a new set of information quantities which we refer to as gradient information.

216, TITLE: SpiderBoost and Momentum: Faster Variance Reduction Algorithms
<https://papers.nips.cc/paper/8511-spiderboost-and-momentum-faster-variance-reduction-algorithms>
AUTHORS: Zhe Wang, Kaiyi Ji, Yi Zhou, Yingbin Liang, Vahid Tarokh

HIGHLIGHT: In this paper, we propose SpiderBoost as an improved scheme, which allows to use a much larger constant-level stepsize while maintaining the same near-optimal oracle complexity, and can be extended with proximal mapping to handle composite optimization (which is nonsmooth and nonconvex) with provable convergence guarantee.

217, **TITLE:** Minimax Optimal Estimation of Approximate Differential Privacy on Neighboring Databases
<https://papers.nips.cc/paper/8512-minimax-optimal-estimation-of-approximate-differential-privacy-on-neighboring-databases>

AUTHORS: Xiyang Liu, Sewoong Oh

HIGHLIGHT: We pose it as a property estimation problem, and study the fundamental trade-offs involved in the accuracy in estimated privacy guarantees and the number of samples required.

218, **TITLE:** Backprop with Approximate Activations for Memory-efficient Network Training
<https://papers.nips.cc/paper/8513-backprop-with-approximate-activations-for-memory-efficient-network-training>

AUTHORS: Ayan Chakrabarti, Benjamin Moseley

HIGHLIGHT: In this paper, we propose a new implementation for back-propagation that significantly reduces memory usage, by enabling the use of approximations with negligible computational cost and minimal effect on training performance.

219, **TITLE:** Training Image Estimators without Image Ground Truth

<https://papers.nips.cc/paper/8514-training-image-estimators-without-image-ground-truth>

AUTHORS: Zhihao Xia, Ayan Chakrabarti

HIGHLIGHT: In this paper, we introduce an unsupervised framework for training image estimation networks, from a training set that contains only measurements---with two varied measurements per image---but no ground-truth for the full images desired as output.

220, **TITLE:** Deep Structured Prediction for Facial Landmark Detection

<https://papers.nips.cc/paper/8515-deep-structured-prediction-for-facial-landmark-detection>

AUTHORS: Lisha Chen, Hui Su, Qiang Ji

HIGHLIGHT: This paper proposes a method for deep structured facial landmark detection based on combining a deep Convolutional Network with a Conditional Random Field.

221, **TITLE:** Information-Theoretic Confidence Bounds for Reinforcement Learning

<https://papers.nips.cc/paper/8516-information-theoretic-confidence-bounds-for-reinforcement-learning>

AUTHORS: Xiuyuan Lu, Benjamin Van Roy

HIGHLIGHT: We integrate information-theoretic concepts into the design and analysis of optimistic algorithms and Thompson sampling.

222, **TITLE:** Transfer Anomaly Detection by Inferring Latent Domain Representations

<https://papers.nips.cc/paper/8517-transfer-anomaly-detection-by-inferring-latent-domain-representations>

AUTHORS: Atsutoshi Kumagai, Tomoharu Iwata, Yasuhiro Fujiwara

HIGHLIGHT: We propose a method to improve the anomaly detection performance on target domains by transferring knowledge on related domains.

223, **TITLE:** Total Least Squares Regression in Input Sparsity Time

<https://papers.nips.cc/paper/8518-total-least-squares-regression-in-input-sparsity-time>

AUTHORS: Huaian Diao, Zhao Song, David Woodruff, Xin Yang

HIGHLIGHT: We give an algorithm for finding a solution \mathbf{X} to the linear system $\mathbf{A}\mathbf{X}=\mathbf{B}$ for which the cost $\|\mathbf{A}-\hat{\mathbf{A}}\|_F^2 + \|\mathbf{B}-\hat{\mathbf{B}}\|_F^2$ is at most a multiplicative $(1+\epsilon)$ factor times the optimal cost, up to an additive error ϵ that may be an arbitrarily small function of n .

224, **TITLE:** Park: An Open Platform for Learning-Augmented Computer Systems

<https://papers.nips.cc/paper/8519-park-an-open-platform-for-learning-augmented-computer-systems>

AUTHORS: Hongzi Mao, Parimarjan Negi, Akshay Narayan, Hanrui Wang, Jiacheng Yang, Haonan Wang, Ryan Marcus, ravichandra addanki, Mehrdad Khani Shirkoohi, Songtao He, Vikram Nathan, Frank Cangialosi, Shailesh Venkatakrisnan, Wei-Hung Weng, Song Han, Tim Kraska, Dr.Mohammad Alizadeh

HIGHLIGHT: We present Park, a platform for researchers to experiment with Reinforcement Learning (RL) for computer systems.

225, **TITLE:** Adapting Neural Networks for the Estimation of Treatment Effects

<https://papers.nips.cc/paper/8520-adapting-neural-networks-for-the-estimation-of-treatment-effects>

AUTHORS: Claudia Shi, David Blei, Victor Veitch

HIGHLIGHT: We propose two adaptations based on insights from the statistical literature on the estimation of treatment effects.

- 226, TITLE: Learning Transferable Graph Exploration
<https://papers.nips.cc/paper/8521-learning-transferable-graph-exploration>
AUTHORS: Hanjun Dai, Yujia Li, Chenglong Wang, Rishabh Singh, Po-Sen Huang, Pushmeet Kohli
HIGHLIGHT: We propose a 'learning to explore' framework where we learn a policy from a distribution of environments.
- 227, TITLE: Conformal Prediction Under Covariate Shift
<https://papers.nips.cc/paper/8522-conformal-prediction-under-covariate-shift>
AUTHORS: Ryan J. Tibshirani, Rina Foygel Barber, Emmanuel Candes, Aaditya Ramdas
HIGHLIGHT: We extend conformal prediction methodology beyond the case of exchangeable data.
- 228, TITLE: Optimal Analysis of Subset-Selection Based L_p Low-Rank Approximation
https://papers.nips.cc/paper/8523-optimal-analysis-of-subset-selection-based-l_p-low-rank-approximation
AUTHORS: Chen Dan, Hong Wang, Hongyang Zhang, Yuchen Zhou, Pradeep K. Ravikumar
HIGHLIGHT: We show that for the problem of ℓ_p rank- k approximation of any given matrix over $\mathbb{R}^{n \times m}$ and $\mathbb{C}^{n \times m}$, the algorithm of column subset selection enjoys approximation ratio $(k+1)^{1/p}$ for $1 \leq p \leq 2$ and $(k+1)^{1-1/p}$ for $p \geq 2$.
- 229, TITLE: Asymmetric Valleys: Beyond Sharp and Flat Local Minima
<https://papers.nips.cc/paper/8524-asymmetric-valleys-beyond-sharp-and-flat-local-minima>
AUTHORS: Haowei He, Gao Huang, Yang Yuan
HIGHLIGHT: In this paper, we observe that local minima of modern deep networks are more than being flat or sharp.
- 230, TITLE: Positive-Unlabeled Compression on the Cloud
<https://papers.nips.cc/paper/8525-positive-unlabeled-compression-on-the-cloud>
AUTHORS: Yixing Xu, Yunhe Wang, Jia Zeng, Kai Han, Chunjing XU, Dacheng Tao, Chang Xu
HIGHLIGHT: In this paper, we present a novel positive-unlabeled (PU) setting for addressing this problem.
- 231, TITLE: Direct Estimation of Differential Functional Graphical Models
<https://papers.nips.cc/paper/8526-direct-estimation-of-differential-functional-graphical-models>
AUTHORS: Boxin Zhao, Y. Samuel Wang, Mladen Kolar
HIGHLIGHT: We consider the problem of estimating the difference between two functional undirected graphical models with shared structures.
- 232, TITLE: On the Calibration of Multiclass Classification with Rejection
<https://papers.nips.cc/paper/8527-on-the-calibration-of-multiclass-classification-with-rejection>
AUTHORS: Chenri Ni, Nontawat Charoenphakdee, Junya Honda, Masashi Sugiyama
HIGHLIGHT: We propose rejection criteria for more general losses for this approach and guarantee calibration to the Bayes-optimal solution.
- 233, TITLE: Third-Person Visual Imitation Learning via Decoupled Hierarchical Controller
<https://papers.nips.cc/paper/8528-third-person-visual-imitation-learning-via-decoupled-hierarchical-controller>
AUTHORS: Pratyusha Sharma, Deepak Pathak, Abhinav Gupta
HIGHLIGHT: Our central insight is to enforce this structure explicitly during learning by decoupling what to achieve (intended task) from how to perform it (controller).
- 234, TITLE: Stagewise Training Accelerates Convergence of Testing Error Over SGD
<https://papers.nips.cc/paper/8529-stagewise-training-accelerates-convergence-of-testing-error-over-sgd>
AUTHORS: Zhuoning Yuan, Yan Yan, Rong Jin, Tianbao Yang
HIGHLIGHT: This paper provides some theoretical evidence for explaining this faster convergence.
- 235, TITLE: Learning Robust Options by Conditional Value at Risk Optimization
<https://papers.nips.cc/paper/8530-learning-robust-options-by-conditional-value-at-risk-optimization>
AUTHORS: Takuya Hiraoka, Takahisa Imagawa, Tatsuya Mori, Takashi Onishi, Yoshimasa Tsuruoka
HIGHLIGHT: In this paper, we propose a conditional value at risk (CVaR)-based method to learn options that work well in both the average and worst cases.
- 236, TITLE: Non-asymptotic Analysis of Stochastic Methods for Non-Smooth Non-Convex Regularized Problems
<https://papers.nips.cc/paper/8531-non-asymptotic-analysis-of-stochastic-methods-for-non-smooth-non-convex-regularized-problems>

AUTHORS: Yi Xu, Rong Jin, Tianbao Yang
HIGHLIGHT: Our contributions are two-fold: (i) we show that they enjoy the same complexities as their counterparts for solving convex regularized non-convex problems in terms of finding an approximate stationary point; (ii) we develop more practical variants using dynamic mini-batch size instead of a fixed mini-batch size without requiring the target accuracy level of solution.

237, TITLE: On Learning Over-parameterized Neural Networks: A Functional Approximation Perspective
<https://papers.nips.cc/paper/8532-on-learning-over-parameterized-neural-networks-a-functional-approximation-perspective>
AUTHORS: Lili Su, Pengkun Yang
HIGHLIGHT: We consider training over-parameterized two-layer neural networks with Rectified Linear Unit (ReLU) using gradient descent (GD) method.

238, TITLE: Drill-down: Interactive Retrieval of Complex Scenes using Natural Language Queries
<https://papers.nips.cc/paper/8533-drill-down-interactive-retrieval-of-complex-scenes-using-natural-language-queries>
AUTHORS: Fuwen Tan, Paola Cascante-Bonilla, Xiaoxiao Guo, Hui Wu, Song Feng, Vicente Ordonez
HIGHLIGHT: We propose Drill-down, an effective framework for encoding multiple queries with an efficient compact state representation that significantly extends current methods for single-round image retrieval.

239, TITLE: Visual Sequence Learning in Hierarchical Prediction Networks and Primate Visual Cortex
<https://papers.nips.cc/paper/8534-visual-sequence-learning-in-hierarchical-prediction-networks-and-primate-visual-cortex>
AUTHORS: JIELIN QIU, Ge Huang, Tai Sing Lee
HIGHLIGHT: In this paper we developed a computational hierarchical network model to understand the spatiotemporal sequence learning effects observed in the primate visual cortex.

240, TITLE: Dual Variational Generation for Low Shot Heterogeneous Face Recognition
<https://papers.nips.cc/paper/8535-dual-variational-generation-for-low-shot-heterogeneous-face-recognition>
AUTHORS: Chaoyou Fu, Xiang Wu, Yibo Hu, Huaibo Huang, Ran He
HIGHLIGHT: This paper considers HFR as a dual generation problem, and proposes a novel Dual Variational Generation (DVG) framework.

241, TITLE: Discovering Neural Wirings
<https://papers.nips.cc/paper/8536-discovering-neural-wirings>
AUTHORS: Mitchell Wortsman, Ali Farhadi, Mohammad Rastegari
HIGHLIGHT: In this work we propose a method for discovering neural wirings.

242, TITLE: On the Optimality of Perturbations in Stochastic and Adversarial Multi-armed Bandit Problems
<https://papers.nips.cc/paper/8537-on-the-optimality-of-perturbations-in-stochastic-and-adversarial-multi-armed-bandit-problems>
AUTHORS: Baekjin Kim, Ambuj Tewari
HIGHLIGHT: We investigate the optimality of perturbation based algorithms in the stochastic and adversarial multi-armed bandit problems.

243, TITLE: Knowledge Extraction with No Observable Data
<https://papers.nips.cc/paper/8538-knowledge-extraction-with-no-observable-data>
AUTHORS: Jaemin Yoo, Minyong Cho, Taebum Kim, U Kang
HIGHLIGHT: In this work, we propose KegNet (Knowledge Extraction with Generative Networks), a novel approach to extract the knowledge of a trained deep neural network and to generate artificial data points that replace the missing training data in knowledge distillation.

244, TITLE: PAC-Bayes under potentially heavy tails
<https://papers.nips.cc/paper/8539-pac-bayes-under-potentially-heavy-tails>
AUTHORS: Matthew Holland
HIGHLIGHT: We derive PAC-Bayesian learning guarantees for heavy-tailed losses, and obtain a novel optimal Gibbs posterior which enjoys finite-sample excess risk bounds at logarithmic confidence.

245, TITLE: One-Shot Object Detection with Co-Attention and Co-Excitation
<https://papers.nips.cc/paper/8540-one-shot-object-detection-with-co-attention-and-co-excitation>
AUTHORS: Ting-I Hsieh, Yi-Chen Lo, Hwann-Tzong Chen, Tyng-Luh Liu
HIGHLIGHT: This paper aims to tackle the challenging problem of one-shot object detection.

246, TITLE: Quaternion Knowledge Graph Embeddings
<https://papers.nips.cc/paper/8541-quaternion-knowledge-graph-embeddings>

- AUTHORS: SHUAI ZHANG, Yi Tay, Lina Yao, Qi Liu
HIGHLIGHT: In this work, we move beyond the traditional complex-valued representations, introducing more expressive hypercomplex representations to model entities and relations for knowledge graph embeddings.
- 247, TITLE: Glyce: Glyph-vectors for Chinese Character Representations
<https://papers.nips.cc/paper/8542-glyce-glyph-vectors-for-chinese-character-representations>
AUTHORS: Yuxian Meng, Wei Wu, Fei Wang, Xiaoya Li, Ping Nie, Fan Yin, Muyu Li, Qinghong Han, Yuxian Meng, Jiwei Li
HIGHLIGHT: In this paper, we address this gap by presenting Glyce, the glyph-vectors for Chinese character representations.
- 248, TITLE: Turbo Autoencoder: Deep learning based channel codes for point-to-point communication channels
<https://papers.nips.cc/paper/8543-turbo-autoencoder-deep-learning-based-channel-codes-for-point-to-point-communication-channels>
AUTHORS: Yihan Jiang, Hyeji Kim, Himanshu Asnani, Sreeram Kannan, Sewoong Oh, Pramod Viswanath
HIGHLIGHT: In this work, we make significant progress on this problem by designing a fully end-to-end jointly trained neural encoder and decoder, namely, Turbo Autoencoder (TurboAE), with the following contributions: (a) under moderate block lengths, TurboAE approaches state-of-the-art performance under canonical channels; (b) moreover, TurboAE outperforms the state-of-the-art codes under non-canonical settings in terms of reliability.
- 249, TITLE: Heterogeneous Graph Learning for Visual Commonsense Reasoning
<https://papers.nips.cc/paper/8544-heterogeneous-graph-learning-for-visual-commonsense-reasoning>
AUTHORS: Weijiang Yu, Jingwen Zhou, Weihao Yu, Xiaodan Liang, Nong Xiao
HIGHLIGHT: In this paper, we propose a new Heterogeneous Graph Learning (HGL) framework for seamlessly integrating the intra-graph and inter-graph reasoning in order to bridge the vision and language domain.
- 250, TITLE: Probabilistic Watershed: Sampling all spanning forests for seeded segmentation and semi-supervised learning
<https://papers.nips.cc/paper/8545-probabilistic-watershed-sampling-all-spanning-forests-for-seeded-segmentation-and-semi-supervised-learning>
AUTHORS: Enrique Fita Sanmartin, Sebastian Damrich, Fred A. Hamprecht
HIGHLIGHT: We propose instead to consider all possible spanning forests and calculate, for every node, the probability of sampling a forest connecting a certain seed with that node.
- 251, TITLE: Classification-by-Components: Probabilistic Modeling of Reasoning over a Set of Components
<https://papers.nips.cc/paper/8546-classification-by-components-probabilistic-modeling-of-reasoning-over-a-set-of-components>
AUTHORS: Sascha Saralajew, Lars Holdijk, Maike Rees, Ebubekir Asan, Thomas Villmann
HIGHLIGHT: In this work, a network architecture, denoted as Classification-By-Components network (CBC), is proposed.
- 252, TITLE: Identifying Causal Effects via Context-specific Independence Relations
<https://papers.nips.cc/paper/8547-identifying-causal-effects-via-context-specific-independence-relations>
AUTHORS: Santtu Tikka, Antti Hyttinen, Juha Karvanen
HIGHLIGHT: Motivated by this, we design a calculus and an automated search procedure for identifying causal effects in the presence of CSIs.
- 253, TITLE: Bridging Machine Learning and Logical Reasoning by Abductive Learning
<https://papers.nips.cc/paper/8548-bridging-machine-learning-and-logical-reasoning-by-abductive-learning>
AUTHORS: Wang-Zhou Dai, Qiuling Xu, Yang Yu, Zhi-Hua Zhou
HIGHLIGHT: In this paper, we present the abductive learning targeted at unifying the two AI paradigms in a mutually beneficial way, where the machine learning model learns to perceive primitive logic facts from data, while logical reasoning can exploit symbolic domain knowledge and correct the wrongly perceived facts for improving the machine learning models.
- 254, TITLE: Regret Minimization for Reinforcement Learning by Evaluating the Optimal Bias Function
<https://papers.nips.cc/paper/8549-regret-minimization-for-reinforcement-learning-by-evaluating-the-optimal-bias-function>
AUTHORS: Zihan Zhang, Xiangyang Ji
HIGHLIGHT: We present an algorithm based on the \emph{Optimism in the Face of Uncertainty} (OFU) principle which is able to learn Reinforcement Learning (RL) modeled by Markov decision process (MDP) with finite state-action space efficiently.
- 255, TITLE: On the Global Convergence of (Fast) Incremental Expectation Maximization Methods
<https://papers.nips.cc/paper/8550-on-the-global-convergence-of-fast-incremental-expectation-maximization-methods>
AUTHORS: Belhal Karimi, Hoi-To Wai, Eric Moulines, Marc Lavielle
HIGHLIGHT: In this paper, we analyze incremental and stochastic version of the EM algorithm as well as the variance reduced-version of [Chen et al., 2018] in a common unifying framework.

- 256, TITLE: A Linearly Convergent Proximal Gradient Algorithm for Decentralized Optimization
<https://papers.nips.cc/paper/8551-a-linearly-convergent-proximal-gradient-algorithm-for-decentralized-optimization>
AUTHORS: Sulaiman Alghunaim, Kun Yuan, Ali H. Sayed
HIGHLIGHT: This work studies decentralized composite optimization problems with non-smooth regularization terms.
- 257, TITLE: Regularizing Trajectory Optimization with Denoising Autoencoders
<https://papers.nips.cc/paper/8552-regularizing-trajectory-optimization-with-denoising-autoencoders>
AUTHORS: Rinu Boney, Norman Di Palo, Mathias Berglund, Alexander Ilin, Juho Kannala, Antti Rasmus, Harri Valpola
HIGHLIGHT: We propose to regularize trajectory optimization by means of a denoising autoencoder that is trained on the same trajectories as the model of the environment.
- 258, TITLE: Learning Hierarchical Priors in VAEs
<https://papers.nips.cc/paper/8553-learning-hierarchical-priors-in-vaes>
AUTHORS: Alexej Klushyn, Nutan Chen, Richard Kurle, Botond Cseke, Patrick van der Smagt
HIGHLIGHT: We introduce a graph-based interpolation method, which shows that the topology of the learned latent representation corresponds to the topology of the data manifold---and present several examples, where desired properties of latent representation such as smoothness and simple explanatory factors are learned by the prior.
- 259, TITLE: Epsilon-Best-Arm Identification in Pay-Per-Reward Multi-Armed Bandits
<https://papers.nips.cc/paper/8554-epsilon-best-arm-identification-in-pay-per-reward-multi-armed-bandits>
AUTHORS: Sivan Sabato
HIGHLIGHT: We provide an algorithm for this setting, that with a high probability returns an epsilon-best arm, while incurring a cost that depends only linearly on the total expected reward of all arms, and does not depend at all on the number of arms.
- 260, TITLE: Safe Exploration for Interactive Machine Learning
<https://papers.nips.cc/paper/8555-safe-exploration-for-interactive-machine-learning>
AUTHORS: Matteo Turchetta, Felix Berkenkamp, Andreas Krause
HIGHLIGHT: In this paper, we introduce a novel framework that renders any existing unsafe IML algorithm safe.
- 261, TITLE: Addressing Failure Detection by Learning Model Confidence
<https://papers.nips.cc/paper/8556-addressing-failure-detection-by-learning-model-confidence>
AUTHORS: Charles Corbiere, Nicolas THOME, Avner Bar-Hen, Matthieu Cord, Patrick P?rez
HIGHLIGHT: In this paper, we propose a new target criterion for model confidence, corresponding to the True Class Probability (TCP).
- 262, TITLE: Combinatorial Bayesian Optimization using the Graph Cartesian Product
<https://papers.nips.cc/paper/8557-combinatorial-bayesian-optimization-using-the-graph-cartesian-product>
AUTHORS: Changyong Oh, Jakub Tomczak, Efstratios Gavves, Max Welling
HIGHLIGHT: We introduce COMBO, a new Gaussian Process (GP) BO.
- 263, TITLE: Fooling Neural Network Interpretations via Adversarial Model Manipulation
<https://papers.nips.cc/paper/8558-fooling-neural-network-interpretations-via-adversarial-model-manipulation>
AUTHORS: Juyeon Heo, Sunghwan Joo, Taesup Moon
HIGHLIGHT: We propose two types of fooling, Passive and Active, and demonstrate such foolings generalize well to the entire validation set as well as transfer to other interpretation methods.
- 264, TITLE: On Lazy Training in Differentiable Programming
<https://papers.nips.cc/paper/8559-on-lazy-training-in-differentiable-programming>
AUTHORS: L?na?c Chizat, Edouard Oyallon, Francis Bach
HIGHLIGHT: In this work, we show that this "lazy training" phenomenon is not specific to over-parameterized neural networks, and is due to a choice of scaling, often implicit, that makes the model behave as its linearization around the initialization, thus yielding a model equivalent to learning with positive-definite kernels.
- 265, TITLE: Quality Aware Generative Adversarial Networks
<https://papers.nips.cc/paper/8560-quality-aware-generative-adversarial-networks>
AUTHORS: KANCHARLA PARIMALA, Sumohana Channappayya
HIGHLIGHT: In this work, we show how a distance metric that is a variant of the Structural SIMilarity (SSIM) index (a popular full-reference image quality assessment algorithm), and a novel quality aware discriminator gradient penalty function that is inspired by the Natural Image Quality Evaluator (NIQE, a popular no-reference image quality assessment algorithm) can each be used as excellent regularizers for GAN objective functions.

- 266, TITLE: Copula-like Variational Inference
<https://papers.nips.cc/paper/8561-copula-like-variational-inference>
AUTHORS: Marcel Hirt, Petros Dellaportas, Alain Durmus
HIGHLIGHT: This paper considers a new family of variational distributions motivated by Sklar's theorem.
- 267, TITLE: Implicit Regularization for Optimal Sparse Recovery
<https://papers.nips.cc/paper/8562-implicit-regularization-for-optimal-sparse-recovery>
AUTHORS: Tomas Vaskevicius, Varun Kanade, Patrick Rebeschini
HIGHLIGHT: We investigate implicit regularization schemes for gradient descent methods applied to unpenalized least squares regression to solve the problem of reconstructing a sparse signal from an underdetermined system of linear measurements under the restricted isometry assumption.
- 268, TITLE: Locally Private Gaussian Estimation
<https://papers.nips.cc/paper/8563-locally-private-gaussian-estimation>
AUTHORS: Matthew Joseph, Janardhan Kulkarni, Jieming Mao, Steven Z. Wu
HIGHLIGHT: We study a basic private estimation problem: each of n users draws a single i.i.d. sample from an unknown Gaussian distribution $N(\mu, \sigma^2)$, and the goal is to estimate μ while guaranteeing local differential privacy for each user.
- 269, TITLE: Multi-mapping Image-to-Image Translation via Learning Disentanglement
<https://papers.nips.cc/paper/8564-multi-mapping-image-to-image-translation-via-learning-disentanglement>
AUTHORS: Xiaoming Yu, Yuanqi Chen, Shan Liu, Thomas Li, Ge Li
HIGHLIGHT: To address this issue, we propose a novel unified model, which bridges these two objectives.
- 270, TITLE: Spatially Aggregated Gaussian Processes with Multivariate Areal Outputs
<https://papers.nips.cc/paper/8565-spatially-aggregated-gaussian-processes-with-multivariate-areal-outputs>
AUTHORS: Yusuke Tanaka, Toshiyuki Tanaka, Tomoharu Iwata, Takeshi Kurashima, Maya Okawa, Yasunori Akagi, Hiroyuki Toda
HIGHLIGHT: We propose a probabilistic model for inferring the multivariate function from multiple areal data sets with various granularities.
- 271, TITLE: Fast Structured Decoding for Sequence Models
<https://papers.nips.cc/paper/8566-fast-structured-decoding-for-sequence-models>
AUTHORS: Zhiqing Sun, Zhuohan Li, Haoqing Wang, Di He, Zi Lin, Zhihong Deng
HIGHLIGHT: Specifically, we design an efficient approximation for Conditional Random Fields (CRF) for non-autoregressive sequence models, and further propose a dynamic transition technique to model positional contexts in the CRF.
- 272, TITLE: Learning Temporal Pose Estimation from Sparsely-Labeled Videos
<https://papers.nips.cc/paper/8567-learning-temporal-pose-estimation-from-sparsely-labeled-videos>
AUTHORS: Gedas Bertasius, Christoph Feichtenhofer, Du Tran, Jianbo Shi, Lorenzo Torresani
HIGHLIGHT: To reduce the need for dense annotations, we propose a PoseWarper network that leverages training videos with sparse annotations (every k frames) to learn to perform dense temporal pose propagation and estimation.
- 273, TITLE: Putting An End to End-to-End: Gradient-Isolated Learning of Representations
<https://papers.nips.cc/paper/8568-putting-an-end-to-end-to-end-gradient-isolated-learning-of-representations>
AUTHORS: Sindy L'we, Peter O'Connor, Bastiaan Veeling
HIGHLIGHT: We propose a novel deep learning method for local self-supervised representation learning that does not require labels nor end-to-end backpropagation but exploits the natural order in data instead.
- 274, TITLE: Scalable Gromov-Wasserstein Learning for Graph Partitioning and Matching
<https://papers.nips.cc/paper/8569-scalable-gromov-wasserstein-learning-for-graph-partitioning-and-matching>
AUTHORS: Hongteng Xu, Dixin Luo, Lawrence Carin
HIGHLIGHT: We propose a scalable Gromov-Wasserstein learning (S-GWL) method and establish a novel and theoretically-supported paradigm for large-scale graph analysis.
- 275, TITLE: Meta-Reinforced Synthetic Data for One-Shot Fine-Grained Visual Recognition
<https://papers.nips.cc/paper/8570-meta-reinforced-synthetic-data-for-one-shot-fine-grained-visual-recognition>
AUTHORS: Satoshi Tsutsui, Yanwei Fu, David Crandall
HIGHLIGHT: To this end, this paper proposes a meta-learning framework to reinforce the generated images by original images so that these images can facilitate one-shot learning.

276, TITLE: Real-Time Reinforcement Learning
<https://papers.nips.cc/paper/8571-real-time-reinforcement-learning>
AUTHORS: Simon Ramstedt, Chris Pal
HIGHLIGHT: In this paper, we introduce a new framework, in which states and actions evolve simultaneously and show how it is related to the classical MDP formulation.

277, TITLE: Robust Multi-agent Counterfactual Prediction
<https://papers.nips.cc/paper/8572-robust-multi-agent-counterfactual-prediction>
AUTHORS: Alexander Peysakhovich, Christian Kroer, Adam Lerer
HIGHLIGHT: We propose a method for analyzing the sensitivity of counterfactual conclusions to violations of these assumptions, which we call robust multi-agent counterfactual prediction (RMAC).

278, TITLE: Approximate Inference Turns Deep Networks into Gaussian Processes
<https://papers.nips.cc/paper/8573-approximate-inference-turns-deep-networks-into-gaussian-processes>
AUTHORS: Mohammad Emtyaz E. Khan, Alexander Immer, Ehsan Abedi, Maciej Korzepa
HIGHLIGHT: In this paper, we show that certain Gaussian posterior approximations for Bayesian DNNs are equivalent to GP posteriors.

279, TITLE: Deep Signature Transforms
<https://papers.nips.cc/paper/8574-deep-signature-transforms>
AUTHORS: Patrick Kidger, Patric Bonnier, Imanol Perez Arribas, Cristopher Salvi, Terry Lyons
HIGHLIGHT: We propose a novel approach which combines the advantages of the signature transform with modern deep learning frameworks.

280, TITLE: Individual Regret in Cooperative Nonstochastic Multi-Armed Bandits
<https://papers.nips.cc/paper/8575-individual-regret-in-cooperative-nonstochastic-multi-armed-bandits>
AUTHORS: Yogev Bar-On, Yishay Mansour
HIGHLIGHT: We present algorithms both for the case that the communication graph is known to all the agents, and for the case that the graph is unknown.

281, TITLE: Convergent Policy Optimization for Safe Reinforcement Learning
<https://papers.nips.cc/paper/8576-convergent-policy-optimization-for-safe-reinforcement-learning>
AUTHORS: Ming Yu, Zhuoran Yang, Mladen Kolar, Zhaoran Wang
HIGHLIGHT: We study the safe reinforcement learning problem with nonlinear function approximation, where policy optimization is formulated as a constrained optimization problem with both the objective and the constraint being nonconvex functions.

282, TITLE: Augmented Neural ODEs
<https://papers.nips.cc/paper/8577-augmented-neural-odes>
AUTHORS: Emilien Dupont, Arnaud Doucet, Yee Whye Teh
HIGHLIGHT: To address these limitations, we introduce Augmented Neural ODEs which, in addition to being more expressive models, are empirically more stable, generalize better and have a lower computational cost than Neural ODEs.

283, TITLE: Thompson Sampling for Multinomial Logit Contextual Bandits
<https://papers.nips.cc/paper/8578-thompson-sampling-for-multinomial-logit-contextual-bandits>
AUTHORS: Min-hwan Oh, Garud Iyengar
HIGHLIGHT: The distinguishing feature in this work is that this feedback has a multinomial logistic distribution.

284, TITLE: Backpropagation-Friendly Eigendecomposition
<https://papers.nips.cc/paper/8579-backpropagation-friendly-eigendecomposition>
AUTHORS: Wei Wang, Zheng Dang, Yinlin Hu, Pascal Fua, Mathieu Salzmann
HIGHLIGHT: In this paper, we introduce a numerically stable and differentiable approach to leveraging eigenvectors in deep networks.

285, TITLE: FastSpeech: Fast, Robust and Controllable Text to Speech
<https://papers.nips.cc/paper/8580-fastspeech-fast-robust-and-controllable-text-to-speech>
AUTHORS: Yi Ren, Yangjun Ruan, Xu Tan, Tao Qin, Sheng Zhao, Zhou Zhao, Tie-Yan Liu
HIGHLIGHT: In this work, we propose a novel feed-forward network based on Transformer to generate mel-spectrogram in parallel for TTS.

- 286, TITLE: Ultrametric Fitting by Gradient Descent
<https://papers.nips.cc/paper/8581-ultrametric-fitting-by-gradient-descent>
AUTHORS: Giovanni Chierchia, Benjamin Perret
HIGHLIGHT: We aim to overcome this limitation by presenting a general optimization framework for ultrametric fitting.
- 287, TITLE: Distinguishing Distributions When Samples Are Strategically Transformed
<https://papers.nips.cc/paper/8582-distinguishing-distributions-when-samples-are-strategically-transformed>
AUTHORS: Hanrui Zhang, Yu Cheng, Vincent Conitzer
HIGHLIGHT: In this paper, we give necessary and sufficient conditions for when the principal can distinguish between agents of "good" and "bad" types, when the type affects the distribution of samples that the agent has access to.
- 288, TITLE: Implicit Regularization of Discrete Gradient Dynamics in Linear Neural Networks
<https://papers.nips.cc/paper/8583-implicit-regularization-of-discrete-gradient-dynamics-in-linear-neural-networks>
AUTHORS: Gauthier Gidel, Francis Bach, Simon Lacoste-Julien
HIGHLIGHT: Using a time rescaling, we show that, with a vanishing initialization and a small enough step size, this dynamics sequentially learns the solutions of a reduced-rank regression with a gradually increasing rank.
- 289, TITLE: Deep Set Prediction Networks
<https://papers.nips.cc/paper/8584-deep-set-prediction-networks>
AUTHORS: Yan Zhang, Jonathon Hare, Adam Prugel-Bennett
HIGHLIGHT: We propose a general model for predicting sets that properly respects the structure of sets and avoids this problem.
- 290, TITLE: DppNet: Approximating Determinantal Point Processes with Deep Networks
<https://papers.nips.cc/paper/8585-dppnet-approximating-determinantal-point-processes-with-deep-networks>
AUTHORS: Zelda E. Mariet, Yaniv Ovadia, Jasper Snoek
HIGHLIGHT: We approach this problem by introducing DppNets: generative deep models that produce DPP-like samples for arbitrary ground sets.
- 291, TITLE: Efficient Communication in Multi-Agent Reinforcement Learning via Variance Based Control
<https://papers.nips.cc/paper/8586-efficient-communication-in-multi-agent-reinforcement-learning-via-variance-based-control>
AUTHORS: Sai Qian Zhang, Qi Zhang, Jieyu Lin
HIGHLIGHT: In this work, we propose Variance Based Control (VBC), a simple yet efficient technique to improve communication efficiency in MARL.
- 292, TITLE: Neural Lyapunov Control
<https://papers.nips.cc/paper/8587-neural-lyapunov-control>
AUTHORS: Ya-Chien Chang, Nima Roohi, Sicun Gao
HIGHLIGHT: We propose new methods for learning control policies and neural network Lyapunov functions for nonlinear control problems, with provable guarantee of stability.
- 293, TITLE: Fully Dynamic Consistent Facility Location
<https://papers.nips.cc/paper/8588-fully-dynamic-consistent-facility-location>
AUTHORS: Vincent Cohen-Addad, Niklas Oskar D. Hjuler, Nikos Parotsidis, David Saulpic, Chris Schwiegelshohn
HIGHLIGHT: In this paper, we focus on general metric spaces and mainly on the facility location problem.
- 294, TITLE: SuperGLUE: A Stickier Benchmark for General-Purpose Language Understanding Systems
<https://papers.nips.cc/paper/8589-superglue-a-stickier-benchmark-for-general-purpose-language-understanding-systems>
AUTHORS: Alex Wang, Yada Pruksachatkun, Nikita Nangia, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, Samuel Bowman
HIGHLIGHT: In this paper we present SuperGLUE, a new benchmark styled after GLUE with a new set of more difficult language understanding tasks, a software toolkit, and a public leaderboard.
- 295, TITLE: A Flexible Generative Framework for Graph-based Semi-supervised Learning
<https://papers.nips.cc/paper/8590-a-flexible-generative-framework-for-graph-based-semi-supervised-learning>
AUTHORS: Jiaqi Ma, Weijing Tang, Ji Zhu, Qiaozhu Mei
HIGHLIGHT: In this work, we propose a flexible generative framework for graph-based semi-supervised learning, which approaches the joint distribution of the node features, labels, and the graph structure.

- 296, TITLE: Inherent Weight Normalization in Stochastic Neural Networks
<https://papers.nips.cc/paper/8591-inherent-weight-normalization-in-stochastic-neural-networks>
AUTHORS: Georgios Detorakis, Sourav Dutta, Abhishek Khanna, Matthew Jerry, Suman Datta, Emre Neftci
HIGHLIGHT: Here, we further demonstrate that always-on multiplicative stochasticity combined with simple threshold neurons provide a sufficient substrate for deep learning machines.
- 297, TITLE: Optimal Decision Tree with Noisy Outcomes
<https://papers.nips.cc/paper/8592-optimal-decision-tree-with-noisy-outcomes>
AUTHORS: Su Jia, viswanath nagarajan, Fatemeh Navidi, R Ravi
HIGHLIGHT: We design new approximation algorithms for both the non-adaptive setting, where the test sequence must be fixed a-priori, and the adaptive setting where the test sequence depends on the outcomes of prior tests.
- 298, TITLE: Meta-Curvature
<https://papers.nips.cc/paper/8593-meta-curvature>
AUTHORS: Eunbyung Park, Junier B. Oliva
HIGHLIGHT: We propose meta-curvature (MC), a framework to learn curvature information for better generalization and fast model adaptation.
- 299, TITLE: Intrinsically Efficient, Stable, and Bounded Off-Policy Evaluation for Reinforcement Learning
<https://papers.nips.cc/paper/8594-intrinsically-efficient-stable-and-bounded-off-policy-evaluation-for-reinforcement-learning>
AUTHORS: Nathan Kallus, Masatoshi Uehara
HIGHLIGHT: We propose new estimators for OPE based on empirical likelihood that are always more efficient than IS, SNIS, and DR and satisfy the same stability and boundedness properties as SNIS.
- 300, TITLE: KerGM: Kernelized Graph Matching
<https://papers.nips.cc/paper/8595-kergm-kernelized-graph-matching>
AUTHORS: Zhen Zhang, Yijian Xiang, Lingfei Wu, Bing Xue, Arye Nehorai
HIGHLIGHT: In our paper, we provide a unifying view for these two problems by introducing new rules for array operations in Hilbert spaces.
- 301, TITLE: Transfusion: Understanding Transfer Learning for Medical Imaging
<https://papers.nips.cc/paper/8596-transfusion-understanding-transfer-learning-for-medical-imaging>
AUTHORS: Maithra Raghu, Chiyuan Zhang, Jon Kleinberg, Samy Bengio
HIGHLIGHT: In this paper, we explore properties of transfer learning for medical imaging.
- 302, TITLE: Adversarial training for free!
<https://papers.nips.cc/paper/8597-adversarial-training-for-free>
AUTHORS: Ali Shafahi, Mahyar Najibi, Mohammad Amin Ghiasi, Zheng Xu, John Dickerson, Christoph Studer, Larry S. Davis, Gavin Taylor, Tom Goldstein
HIGHLIGHT: We present an algorithm that eliminates the overhead cost of generating adversarial examples by recycling the gradient information computed when updating model parameters.
- 303, TITLE: Communication-Efficient Distributed Learning via Lazily Aggregated Quantized Gradients
<https://papers.nips.cc/paper/8598-communication-efficient-distributed-learning-via-lazily-aggregated-quantized-gradients>
AUTHORS: Jun Sun, Tianyi Chen, Georgios Giannakis, Zaiyue Yang
HIGHLIGHT: The present paper develops a novel aggregated gradient approach for distributed machine learning that adaptively compresses the gradient communication.
- 304, TITLE: Implicitly learning to reason in first-order logic
<https://papers.nips.cc/paper/8599-implicitly-learning-to-reason-in-first-order-logic>
AUTHORS: Vaishak Belle, Brendan Juba
HIGHLIGHT: In this work, we present a new theoretical approach to robustly learning to reason in first-order logic, and consider universally quantified clauses over a countably infinite domain.
- 305, TITLE: Kernel-Based Approaches for Sequence Modeling: Connections to Neural Methods
<https://papers.nips.cc/paper/8600-kernel-based-approaches-for-sequence-modeling-connections-to-neural-methods>
AUTHORS: Kevin Liang, Guoyin Wang, Yitong Li, Ricardo Henao, Lawrence Carin
HIGHLIGHT: We investigate time-dependent data analysis from the perspective of recurrent kernel machines, from which models with hidden units and gated memory cells arise naturally.

- 306, TITLE: PC-Fairness: A Unified Framework for Measuring Causality-based Fairness
<https://papers.nips.cc/paper/8601-pc-fairness-a-unified-framework-for-measuring-causality-based-fairness>
AUTHORS: Yongkai Wu, Lu Zhang, Xintao Wu, Hanghang Tong
HIGHLIGHT: In this paper, we develop a framework for measuring different causality-based fairness.
- 307, TITLE: Arbicon-Net: Arbitrary Continuous Geometric Transformation Networks for Image Registration
<https://papers.nips.cc/paper/8602-arbicon-net-arbitrary-continuous-geometric-transformation-networks-for-image-registration>
AUTHORS: Jianchun Chen, Lingjing Wang, Xiang Li, Yi Fang
HIGHLIGHT: To address this issue, we present an end-to-end trainable deep neural networks, named Arbitrary Continuous Geometric Transformation Networks (Arbicon-Net), to directly predict the dense displacement field for pairwise image alignment.
- 308, TITLE: Assessing Disparate Impact of Personalized Interventions: Identifiability and Bounds
<https://papers.nips.cc/paper/8603-assessing-disparate-impact-of-personalized-interventions-identifiability-and-bounds>
AUTHORS: Nathan Kallus, Angela Zhou
HIGHLIGHT: We prove how we can nonetheless point-identify these quantities under the additional assumption of monotone treatment response, which may be reasonable in many applications.
- 309, TITLE: The Fairness of Risk Scores Beyond Classification: Bipartite Ranking and the XAUC Metric
<https://papers.nips.cc/paper/8604-the-fairness-of-risk-scores-beyond-classification-bipartite-ranking-and-the-xauc-metric>
AUTHORS: Nathan Kallus, Angela Zhou
HIGHLIGHT: To better account for this, in this paper, we investigate the fairness of predictive risk scores from the point of view of a bipartite ranking task, where one seeks to rank positive examples higher than negative ones.
- 310, TITLE: HYPE: A Benchmark for Human eYe Perceptual Evaluation of Generative Models
<https://papers.nips.cc/paper/8605-hype-a-benchmark-for-human-eye-perceptual-evaluation-of-generative-models>
AUTHORS: Sharon Zhou, Mitchell Gordon, Ranjay Krishna, Austin Narcomey, Li F. Fei-Fei, Michael Bernstein
HIGHLIGHT: Our work establishes a gold standard human benchmark for generative realism.
- 311, TITLE: First order expansion of convex regularized estimators
<https://papers.nips.cc/paper/8606-first-order-expansion-of-convex-regularized-estimators>
AUTHORS: Pierre Bellec, Arun Kuchibhotla
HIGHLIGHT: We consider first order expansions of convex penalized estimators in high-dimensional regression problems with random designs.
- 312, TITLE: Capacity Bounded Differential Privacy
<https://papers.nips.cc/paper/8607-capacity-bounded-differential-privacy>
AUTHORS: Kamalika Chaudhuri, Jacob Imola, Ashwin Machanavajjhala
HIGHLIGHT: In this work, we present a novel relaxation of differential privacy, capacity bounded differential privacy, where the adversary that distinguishes output distributions is assumed to be capacity-bounded -- i.e. bounded not in computational power, but in terms of the function class from which their attack algorithm is drawn.
- 313, TITLE: Universal Boosting Variational Inference
<https://papers.nips.cc/paper/8608-universal-boosting-variational-inference>
AUTHORS: Trevor Campbell, Xinglong Li
HIGHLIGHT: We thus develop universal boosting variational inference (UBVI), a BVI scheme that exploits the simple geometry of probability densities under the Hellinger metric to prevent the degeneracy of other gradient-based BVI methods, avoid difficult joint optimizations of both component and weight, and simplify fully-corrective weight optimizations.
- 314, TITLE: SGD on Neural Networks Learns Functions of Increasing Complexity
<https://papers.nips.cc/paper/8609-sgd-on-neural-networks-learns-functions-of-increasing-complexity>
AUTHORS: Dimitris Kalimeris, Gal Kaplun, Preetum Nakkiran, Benjamin Edelman, Tristan Yang, Boaz Barak, Haofeng Zhang
HIGHLIGHT: More generally, we give evidence for the hypothesis that, as iterations progress, SGD learns functions of increasing complexity.
- 315, TITLE: The Landscape of Non-convex Empirical Risk with Degenerate Population Risk
<https://papers.nips.cc/paper/8610-the-landscape-of-non-convex-empirical-risk-with-degenerate-population-risk>
AUTHORS: Shuang Li, Gongguo Tang, Michael B. Wakin
HIGHLIGHT: In this work, we focus on the situation where the corresponding population risk is a degenerate non-convex loss function, namely, the Hessian of the population risk can have zero eigenvalues.

- 316, TITLE: Making AI Forget You: Data Deletion in Machine Learning
<https://papers.nips.cc/paper/8611-making-ai-forget-you-data-deletion-in-machine-learning>
AUTHORS: Antonio Ginart, Melody Guan, Gregory Valiant, James Y. Zou
HIGHLIGHT: In this paper we initiate a framework studying what to do when it is no longer permissible to deploy models derivative from specific user data.
- 317, TITLE: Practical Differentially Private Top-k Selection with Pay-what-you-get Composition
<https://papers.nips.cc/paper/8612-practical-differentially-private-top-k-selection-with-pay-what-you-get-composition>
AUTHORS: David Durfee, Ryan M. Rogers
HIGHLIGHT: We study the problem of top-k selection over a large domain universe subject to user-level differential privacy.
- 318, TITLE: Conformalized Quantile Regression
<https://papers.nips.cc/paper/8613-conformalized-quantile-regression>
AUTHORS: Yaniv Romano, Evan Patterson, Emmanuel Candes
HIGHLIGHT: In this paper we propose a new method that is fully adaptive to heteroscedasticity.
- 319, TITLE: Thompson Sampling with Information Relaxation Penalties
<https://papers.nips.cc/paper/8614-thompson-sampling-with-information-relaxation-penalties>
AUTHORS: Seungki Min, Costis Maglaras, Ciamac C. Moallemi
HIGHLIGHT: We consider a finite-horizon multi-armed bandit (MAB) problem in a Bayesian setting, for which we propose an information relaxation sampling framework.
- 320, TITLE: Deep Generalized Method of Moments for Instrumental Variable Analysis
<https://papers.nips.cc/paper/8615-deep-generalized-method-of-moments-for-instrumental-variable-analysis>
AUTHORS: Andrew Bennett, Nathan Kallus, Tobias Schnabel
HIGHLIGHT: In this paper, we propose the DeepGMM algorithm to overcome this.
- 321, TITLE: Learning Sample-Specific Models with Low-Rank Personalized Regression
<https://papers.nips.cc/paper/8616-learning-sample-specific-models-with-low-rank-personalized-regression>
AUTHORS: Ben Lengerich, Bryon Aragam, Eric P. Xing
HIGHLIGHT: To address this challenge, we propose to estimate sample-specific models that tailor inference and prediction at the individual level.
- 322, TITLE: Dancing to Music
<https://papers.nips.cc/paper/8617-dancing-to-music>
AUTHORS: Hsin-Ying Lee, Xiaodong Yang, Ming-Yu Liu, Ting-Chun Wang, Yu-Ding Lu, Ming-Hsuan Yang, Jan Kautz
HIGHLIGHT: In this paper, we propose a synthesis-by-analysis learning framework to generate dance from music.
- 323, TITLE: Deconstructing Lottery Tickets: Zeros, Signs, and the Supermask
<https://papers.nips.cc/paper/8618-deconstructing-lottery-tickets-zeros-signs-and-the-supermask>
AUTHORS: Hattie Zhou, Janice Lan, Rosanne Liu, Jason Yosinski
HIGHLIGHT: In this paper we study the three critical components of the Lottery Ticket (LT) algorithm, showing that each may be varied significantly without impacting the overall results.
- 324, TITLE: Implicit Generation and Modeling with Energy Based Models
<https://papers.nips.cc/paper/8619-implicit-generation-and-modeling-with-energy-based-models>
AUTHORS: Yilun Du, Igor Mordatch
HIGHLIGHT: We present techniques to scale MCMC based EBM training on continuous neural networks, and we show its success on the high-dimensional data domains of ImageNet32x32, ImageNet128x128, CIFAR-10, and robotic hand trajectories, achieving better samples than other likelihood models and nearing the performance of contemporary GAN approaches, while covering all modes of the data.
- 325, TITLE: LCA: Loss Change Allocation for Neural Network Training
<https://papers.nips.cc/paper/8620-lca-loss-change-allocation-for-neural-network-training>
AUTHORS: Janice Lan, Rosanne Liu, Hattie Zhou, Jason Yosinski
HIGHLIGHT: We propose a new window into training called Loss Change Allocation (LCA), in which credit for changes to the network loss is conservatively partitioned to the parameters.

- 326, TITLE: Predicting the Politics of an Image Using Webly Supervised Data
<https://papers.nips.cc/paper/8621-predicting-the-politics-of-an-image-using-webly-supervised-data>
AUTHORS: Christopher Thomas, Adriana Kovashka
HIGHLIGHT: In this paper, we model visual political bias in contemporary media sources at scale, using webly supervised data.
We collect a dataset of over one million unique images and associated news articles from left- and right-leaning news sources, and develop a method to predict the image's political leaning.
- 327, TITLE: Adaptive GNN for Image Analysis and Editing
<https://papers.nips.cc/paper/8622-adaptive-gnn-for-image-analysis-and-editing>
AUTHORS: Lingyu Liang, LianWen Jin, Yong Xu
HIGHLIGHT: In mathematical analysis, we propose an adaptive GNN model by recursive definition, and derive its relation with two basic operations in CV: filtering and propagation operations.
- 328, TITLE: Ultra Fast Medoid Identification via Correlated Sequential Halving
<https://papers.nips.cc/paper/8623-ultra-fast-medoid-identification-via-correlated-sequential-halving>
AUTHORS: Tavor Baharav, David Tse
HIGHLIGHT: In this work, we show that we can better exploit the structure of the underlying computation problem by modifying the traditional bandit sampling strategy and using it in conjunction with a suitably chosen multi-armed bandit algorithm.
- 329, TITLE: Tight Dimension Independent Lower Bound on the Expected Convergence Rate for Diminishing Step Sizes in SGD
<https://papers.nips.cc/paper/8624-tight-dimension-independent-lower-bound-on-the-expected-convergence-rate-for-diminishing-step-sizes-in-sgd>
AUTHORS: PHUONG HA NGUYEN, Lam Nguyen, Marten van Dijk
HIGHLIGHT: We study the convergence of Stochastic Gradient Descent (SGD) for strongly convex objective functions.
- 330, TITLE: Asymptotics for Sketching in Least Squares Regression
<https://papers.nips.cc/paper/8625-asymptotics-for-sketching-in-least-squares-regression>
AUTHORS: Edgar Dobriban, Sifan Liu
HIGHLIGHT: In this paper, we make progress on this problem, working in an asymptotic framework where the number of datapoints and dimension of features goes to infinity.
- 331, TITLE: MCP: Learning Composable Hierarchical Control with Multiplicative Compositional Policies
<https://papers.nips.cc/paper/8626-mcp-learning-composable-hierarchical-control-with-multiplicative-compositional-policies>
AUTHORS: Xue Bin Peng, Michael Chang, Grace Zhang, Pieter Abbeel, Sergey Levine
HIGHLIGHT: In this work, we propose multiplicative compositional policies (MCP), a method for learning reusable motor skills that can be composed to produce a range of complex behaviors.
- 332, TITLE: Exact inference in structured prediction
<https://papers.nips.cc/paper/8627-exact-inference-in-structured-prediction>
AUTHORS: Kevin Bello, Jean Honorio
HIGHLIGHT: We consider the generative process proposed by Globerson et al. (2015) and apply it to general connected graphs.
- 333, TITLE: Coda: An End-to-End Neural Program Decompiler
<https://papers.nips.cc/paper/8628-coda-an-end-to-end-neural-program-decompiler>
AUTHORS: Cheng Fu, Huili Chen, Haolan Liu, Xinyun Chen, Yuandong Tian, Farinaz Koushanfar, Jishen Zhao
HIGHLIGHT: To address the above problems, we propose Coda1, the first end-to-end neural-based framework for code decompilation.
- 334, TITLE: Bat-G net: Bat-inspired High-Resolution 3D Image Reconstruction using Ultrasonic Echoes
<https://papers.nips.cc/paper/8629-bat-g-net-bat-inspired-high-resolution-3d-image-reconstruction-using-ultrasonic-echoes>
AUTHORS: Gunpil Hwang, Seohyeon Kim, Hyeon-Min Bae
HIGHLIGHT: In this paper, a bat-inspired high-resolution ultrasound 3D imaging system is presented.
- 335, TITLE: Painless Stochastic Gradient: Interpolation, Line-Search, and Convergence Rates
<https://papers.nips.cc/paper/8630-painless-stochastic-gradient-interpolation-line-search-and-convergence-rates>
AUTHORS: Sharan Vaswani, Aaron Mishkin, Issam Laradji, Mark Schmidt, Gauthier Gidel, Simon Lacoste-Julien
HIGHLIGHT: We propose to use line-search techniques to automatically set the step-size when training models that can interpolate the data.

- 336, TITLE: Scalable Structure Learning of Continuous-Time Bayesian Networks from Incomplete Data
<https://papers.nips.cc/paper/8631-scalable-structure-learning-of-continuous-time-bayesian-networks-from-incomplete-data>
AUTHORS: Dominik Linzner, Michael Schmidt, Heinz Koepl
HIGHLIGHT: Instead of sampling and scoring all possible structures individually, we assume the generator of the CTBN to be composed as a mixture of generators stemming from different structures. In this framework, structure learning can be performed via a gradient-based optimization of mixture weights.
- 337, TITLE: Privacy-Preserving Classification of Personal Text Messages with Secure Multi-Party Computation
<https://papers.nips.cc/paper/8632-privacy-preserving-classification-of-personal-text-messages-with-secure-multi-party-computation>
AUTHORS: Devin Reich, Ariel Todoki, Rafael Dowsley, Martine De Cock, anderson nascimento
HIGHLIGHT: We propose the first privacy-preserving solution for text classification that is provably secure.
- 338, TITLE: Efficiently Estimating Erdos-Renyi Graphs with Node Differential Privacy
<https://papers.nips.cc/paper/8633-efficiently-estimating-erdos-renyi-graphs-with-node-differential-privacy>
AUTHORS: Jonathan Ullman, Adam Sealfon
HIGHLIGHT: We give a simple, computationally efficient, and node-differentially-private algorithm for estimating the parameter of an Erdos-Renyi graph---that is, estimating p in a $G(n,p)$ ---with near-optimal accuracy.
- 339, TITLE: Learning Representations for Time Series Clustering
<https://papers.nips.cc/paper/8634-learning-representations-for-time-series-clustering>
AUTHORS: Qianli Ma, Jiawei Zheng, Sen Li, Gary W. Cottrell
HIGHLIGHT: Here we propose a novel unsupervised temporal representation learning model, named Deep Temporal Clustering Representation (DTCR), which integrates the temporal reconstruction and K-means objective into the seq2seq model.
- 340, TITLE: Verified Uncertainty Calibration
<https://papers.nips.cc/paper/8635-verified-uncertainty-calibration>
AUTHORS: Ananya Kumar, Percy S. Liang, Tengyu Ma
HIGHLIGHT: To get the best of both worlds, we introduce the scaling-binning calibrator, which first fits a parametric function that acts like a baseline for variance reduction and then bins the function values to actually ensure calibration.
- 341, TITLE: A Normative Theory for Causal Inference and Bayes Factor Computation in Neural Circuits
<https://papers.nips.cc/paper/8636-a-normative-theory-for-causal-inference-and-bayes-factor-computation-in-neural-circuits>
AUTHORS: Wenhao Zhang, Si Wu, Brent Doiron, Tai Sing Lee
HIGHLIGHT: In this paper, we consider the causal inference in multisensory processing and propose a novel generative model based on neural population code that takes into account both stimulus feature and stimulus reliability in the inference.
- 342, TITLE: Unsupervised Keypoint Learning for Guiding Class-Conditional Video Prediction
<https://papers.nips.cc/paper/8637-unsupervised-keypoint-learning-for-guiding-class-conditional-video-prediction>
AUTHORS: Yunji Kim, Seonghyeon Nam, In Cho, Seon Joo Kim
HIGHLIGHT: We propose a deep video prediction model conditioned on a single image and an action class.
- 343, TITLE: Subspace Attack: Exploiting Promising Subspaces for Query-Efficient Black-box Attacks
<https://papers.nips.cc/paper/8638-subspace-attack-exploiting-promising-subspaces-for-query-efficient-black-box-attacks>
AUTHORS: Yiwen Guo, Ziang Yan, Changshui Zhang
HIGHLIGHT: In this paper, we aim at reducing the query complexity of black-box attacks in this category.
- 344, TITLE: Stochastic Gradient Hamiltonian Monte Carlo Methods with Recursive Variance Reduction
<https://papers.nips.cc/paper/8639-stochastic-gradient-hamiltonian-monte-carlo-methods-with-recursive-variance-reduction>
AUTHORS: Difan Zou, Pan Xu, Quanquan Gu
HIGHLIGHT: In this paper, we propose a Stochastic Recursive Variance-Reduced gradient HMC (SRVR-HMC) algorithm.
- 345, TITLE: Learning Latent Process from High-Dimensional Event Sequences via Efficient Sampling
<https://papers.nips.cc/paper/8640-learning-latent-process-from-high-dimensional-event-sequences-via-efficient-sampling>
AUTHORS: Qitian Wu, Zixuan Zhang, Xiaofeng Gao, Junchi Yan, Guihai Chen
HIGHLIGHT: To these ends, in this paper, we propose a seminal adversarial imitation learning framework for high-dimension event sequence generation which could be decomposed into: 1) a latent structural intensity model that estimates the adjacent nodes without explicit networks and learns to capture the temporal dynamics in the latent space of markers over observed sequence; 2) an efficient random walk based generation model that aims at imitating the generation process of high-dimension event sequences from a

bottom-up view; 3) a discriminator specified as a seq2seq network optimizing the rewards to help the generator output event sequences as real as possible.

- 346, TITLE: Cross-sectional Learning of Extremal Dependence among Financial Assets
<https://papers.nips.cc/paper/8641-cross-sectional-learning-of-extremal-dependence-among-financial-assets>
AUTHORS: Xing Yan, Qi Wu, Wen Zhang
HIGHLIGHT: We propose a novel probabilistic model to facilitate the learning of multivariate tail dependence of multiple financial assets.
- 347, TITLE: Principal Component Projection and Regression in Nearly Linear Time through Asymmetric SVRG
<https://papers.nips.cc/paper/8642-principal-component-projection-and-regression-in-nearly-linear-time-through-asymmetric-svrg>
AUTHORS: Yujia Jin, Aaron Sidford
HIGHLIGHT: In this paper we provide the first algorithms that solve these problems in nearly linear time for fixed eigenvalue distribution and large n .
- 348, TITLE: Compression with Flows via Local Bits-Back Coding
<https://papers.nips.cc/paper/8643-compression-with-flows-via-local-bits-back-coding>
AUTHORS: Jonathan Ho, Evan Lohn, Pieter Abbeel
HIGHLIGHT: To fill in this gap, we introduce local bits-back coding, a new compression technique for flow models.
- 349, TITLE: Exact Rate-Distortion in Autoencoders via Echo Noise
<https://papers.nips.cc/paper/8644-exact-rate-distortion-in-autoencoders-via-echo-noise>
AUTHORS: Rob Breckelmanns, Daniel Moyer, Aram Galstyan, Greg Ver Steeg
HIGHLIGHT: We introduce a new noise channel, Echo noise, that admits a simple, exact expression for mutual information for arbitrary input distributions.
- 350, TITLE: iSplit LBI: Individualized Partial Ranking with Ties via Split LBI
<https://papers.nips.cc/paper/8645-isplit-lbi-individualized-partial-ranking-with-ties-via-split-lbi>
AUTHORS: Qianqian Xu, Xinwei Sun, Zhiyong Yang, Xiaochun Cao, Qingming Huang, Yuan Yao
HIGHLIGHT: In this paper, instead of learning a global ranking which is agreed with the consensus, we pursue the tie-aware partial ranking from an individualized perspective.
- 351, TITLE: Domes to Drones: Self-Supervised Active Triangulation for 3D Human Pose Reconstruction
<https://papers.nips.cc/paper/8646-domes-to-drones-self-supervised-active-triangulation-for-3d-human-pose-reconstruction>
AUTHORS: Aleksis Pirinen, Erik G?rtner, Cristian Sminchisescu
HIGHLIGHT: In order to address the view selection problem in a principled way, we here introduce ACTOR, an active triangulation agent for 3d human pose reconstruction.
- 352, TITLE: MetaQuant: Learning to Quantize by Learning to Penetrate Non-differentiable Quantization
<https://papers.nips.cc/paper/8647-metaquant-learning-to-quantize-by-learning-to-penetrate-non-differentiable-quantization>
AUTHORS: Shangyu Chen, Wenya Wang, Sinno Jialin Pan
HIGHLIGHT: In this paper, we propose to learn \mathcal{G}_r by a neural network.
- 353, TITLE: Improved Precision and Recall Metric for Assessing Generative Models
<https://papers.nips.cc/paper/8648-improved-precision-and-recall-metric-for-assessing-generative-models>
AUTHORS: Tuomas Kynk?nniemi, Tero Karras, Samuli Laine, Jaakko Lehtinen, Timo Aila
HIGHLIGHT: We present an evaluation metric that can separately and reliably measure both of these aspects in image generation tasks by forming explicit, non-parametric representations of the manifolds of real and generated data.
- 354, TITLE: A First-Order Algorithmic Framework for Distributionally Robust Logistic Regression
<https://papers.nips.cc/paper/8649-a-first-order-algorithmic-framework-for-distributionally-robust-logistic-regression>
AUTHORS: JIAJIN LI, SEN HUANG, Anthony Man-Cho So
HIGHLIGHT: In this paper, we take a first step towards resolving the above difficulty by developing a first-order algorithmic framework for tackling a class of Wasserstein distance-based distributionally robust logistic regression (DRLR) problem.
- 355, TITLE: PasteGAN: A Semi-Parametric Method to Generate Image from Scene Graph
<https://papers.nips.cc/paper/8650-pastegan-a-semi-parametric-method-to-generate-image-from-scene-graph>
AUTHORS: Yikang LI, Tao Ma, Yeqi Bai, Nan Duan, Sining Wei, Xiaogang Wang

HIGHLIGHT: Therefore, to generate the images with preferred objects and rich interactions, we propose a semi-parametric method, PasteGAN, for generating the image from the scene graph and the image crops, where spatial arrangements of the objects and their pair-wise relationships are defined by the scene graph and the object appearances are determined by the given object crops.

356, **TITLE:** Handling correlated and repeated measurements with the smoothed multivariate square-root Lasso
<https://papers.nips.cc/paper/8651-handling-correlated-and-repeated-measurements-with-the-smoothed-multivariate-square-root-lasso>
AUTHORS: Quentin Bertrand, Mathurin Massias, Alexandre Gramfort, Joseph Salmon
HIGHLIGHT: In this work, we propose a concomitant estimator that can cope with complex noise structure by using non-averaged measurements, its data-fitting term arising as a smoothing of the nuclear norm.

357, **TITLE:** Joint Optimization of Tree-based Index and Deep Model for Recommender Systems
<https://papers.nips.cc/paper/8652-joint-optimization-of-tree-based-index-and-deep-model-for-recommender-systems>
AUTHORS: Han Zhu, Daqing Chang, Ziru Xu, Pengye Zhang, Xiang Li, Jie He, Han Li, Jian Xu, Kun Gai
HIGHLIGHT: Our purpose, in this paper, is to develop a method to jointly learn the index structure and user preference prediction model.

358, **TITLE:** Learning Generalizable Device Placement Algorithms for Distributed Machine Learning
<https://papers.nips.cc/paper/8653-learning-generalizable-device-placement-algorithms-for-distributed-machine-learning>
AUTHORS: ravichandra addanki, Shaileshh Bojja Venkatakrishnan, Shreyan Gupta, Hongzi Mao, Mohammad Alizadeh
HIGHLIGHT: We present Placeto, a reinforcement learning (RL) approach to efficiently find device placements for distributed neural network training.

359, **TITLE:** Uncoupled Regression from Pairwise Comparison Data
<https://papers.nips.cc/paper/8654-uncoupled-regression-from-pairwise-comparison-data>
AUTHORS: Liyuan Xu, Junya Honda, Gang Niu, Masashi Sugiyama
HIGHLIGHT: We propose two practical methods for uncoupled regression from pairwise comparison data and show that the learned regression model converges to the optimal model with the optimal parametric convergence rate when the target variable distributes uniformly.

360, **TITLE:** Cross Attention Network for Few-shot Classification
<https://papers.nips.cc/paper/8655-cross-attention-network-for-few-shot-classification>
AUTHORS: Ruibing Hou, Hong Chang, Bingpeng MA, Shiguang Shan, Xilin Chen
HIGHLIGHT: In this work, we propose a novel Cross Attention Network to address the challenging problems in few-shot classification.

361, **TITLE:** A Nonconvex Approach for Exact and Efficient Multichannel Sparse Blind Deconvolution
<https://papers.nips.cc/paper/8656-a-nonconvex-approach-for-exact-and-efficient-multichannel-sparse-blind-deconvolution>
AUTHORS: Qing Qu, Xiao Li, Zhihui Zhu
HIGHLIGHT: We study the multi-channel sparse blind deconvolution (MCS-BD) problem, whose task is to simultaneously recover a kernel \mathbf{a} and multiple sparse inputs $\{\mathbf{x}_i\}_{i=1}^p$ from their circulant convolution $\mathbf{y}_i = \mathbf{a} \circledast \mathbf{x}_i$ ($i=1, \dots, p$).

362, **TITLE:** SCAN: A Scalable Neural Networks Framework Towards Compact and Efficient Models
<https://papers.nips.cc/paper/8657-scan-a-scalable-neural-networks-framework-towards-compact-and-efficient-models>
AUTHORS: Linfeng Zhang, Zhanhong Tan, Jiebo Song, Jingwei Chen, Chenglong Bao, Kaisheng Ma
HIGHLIGHT: To address this problem, we propose the so-called SCAN framework for networks training and inference, which is orthogonal and complementary to existing acceleration and compression methods.

363, **TITLE:** Revisiting the Bethe-Hessian: Improved Community Detection in Sparse Heterogeneous Graphs
<https://papers.nips.cc/paper/8658-revisiting-the-bethe-hessian-improved-community-detection-in-sparse-heterogeneous-graphs>
AUTHORS: Lorenzo Dall'Amico, Romain Couillet, Nicolas Tremblay
HIGHLIGHT: This article studies spectral clustering based on the Bethe-Hessian matrix $H_r = (r^2 - 1)I_n + D - rA$ for sparse heterogeneous graphs (following the degree-corrected stochastic block model) in a two-class setting.

364, **TITLE:** Teaching Multiple Concepts to a Forgetful Learner
<https://papers.nips.cc/paper/8659-teaching-multiple-concepts-to-a-forgetful-learner>
AUTHORS: Anette Hunziker, Yuxin Chen, Oisín Mac Aodha, Manuel Gomez Rodriguez, Andreas Krause, Pietro Perona, Yisong Yue, Adish Singla
HIGHLIGHT: In this paper, we look at the problem from the perspective of discrete optimization and introduce a novel algorithmic framework for teaching multiple concepts with strong performance guarantees.

- 365, TITLE: Regularized Weighted Low Rank Approximation
<https://papers.nips.cc/paper/8660-regularized-weighted-low-rank-approximation>
AUTHORS: Frank Ban, David Woodruff, Richard Zhang
HIGHLIGHT: We derive provably sharper guarantees for the regularized version by obtaining parameterized complexity bounds in terms of the statistical dimension rather than the rank, allowing for a rank-independent runtime that can be significantly faster.
- 366, TITLE: Practical and Consistent Estimation of f-Divergences
<https://papers.nips.cc/paper/8661-practical-and-consistent-estimation-of-f-divergences>
AUTHORS: Paul Rubenstein, Olivier Bousquet, Josip Djolonga, Carlos Riquelme, Ilya O. Tolstikhin
HIGHLIGHT: Under these assumptions we propose and study an estimator that can be easily implemented, works well in high dimensions, and enjoys faster rates of convergence.
- 367, TITLE: Approximation Ratios of Graph Neural Networks for Combinatorial Problems
<https://papers.nips.cc/paper/8662-approximation-ratios-of-graph-neural-networks-for-combinatorial-problems>
AUTHORS: Ryoma Sato, Makoto Yamada, Hisashi Kashima
HIGHLIGHT: In this paper, from a theoretical perspective, we study how powerful graph neural networks (GNNs) can be for learning approximation algorithms for combinatorial problems.
- 368, TITLE: Thinning for Accelerating the Learning of Point Processes
<https://papers.nips.cc/paper/8663-thinning-for-accelerating-the-learning-of-point-processes>
AUTHORS: Tianbo Li, Yiping Ke
HIGHLIGHT: We propose thinning as a downsampling method for accelerating the learning of point processes.
- 369, TITLE: A Prior of a Googol Gaussians: a Tensor Ring Induced Prior for Generative Models
<https://papers.nips.cc/paper/8664-a-prior-of-a-googol-gaussians-a-tensor-ring-induced-prior-for-generative-models>
AUTHORS: Maxim Kuznetsov, Daniil Polykovskiy, Dmitry P. Vetrov, Alex Zhebrak
HIGHLIGHT: Altogether, we propose a novel plug-and-play framework for generative models that can be utilized in any GAN and VAE-like architectures.
- 370, TITLE: Differentially Private Markov Chain Monte Carlo
<https://papers.nips.cc/paper/8665-differentially-private-markov-chain-monte-carlo>
AUTHORS: Mikko Heikkilä, Joonas Jylk, Onur Dikmen, Antti Honkela
HIGHLIGHT: In this paper, we further extend the applicability of DP Bayesian learning by presenting the first general DP Markov chain Monte Carlo (MCMC) algorithm whose privacy-guarantees are not subject to unrealistic assumptions on Markov chain convergence and that is applicable to posterior inference in arbitrary models.
- 371, TITLE: Full-Gradient Representation for Neural Network Visualization
<https://papers.nips.cc/paper/8666-full-gradient-representation-for-neural-network-visualization>
AUTHORS: Suraj Srinivas, François Fleuret
HIGHLIGHT: We introduce a new tool for interpreting neural nets, namely full-gradients, which decomposes the neural net response into input sensitivity and per-neuron sensitivity components.
- 372, TITLE: q-means: A quantum algorithm for unsupervised machine learning
<https://papers.nips.cc/paper/8667-q-means-a-quantum-algorithm-for-unsupervised-machine-learning>
AUTHORS: Iordanis Kerenidis, Jonas Landman, Alessandro Luongo, Anupam Prakash
HIGHLIGHT: In this paper, we introduce q-means, a new quantum algorithm for clustering.
- 373, TITLE: Learner-aware Teaching: Inverse Reinforcement Learning with Preferences and Constraints
<https://papers.nips.cc/paper/8668-learner-aware-teaching-inverse-reinforcement-learning-with-preferences-and-constraints>
AUTHORS: Sebastian Tschiatschek, Ahana Ghosh, Luis Haug, Rati Devidze, Adish Singla
HIGHLIGHT: In this paper, we consider the setting where the learner has its own preferences that it additionally takes into consideration.
- 374, TITLE: Limitations of the empirical Fisher approximation for natural gradient descent
<https://papers.nips.cc/paper/8669-limitations-of-the-empirical-fisher-approximation-for-natural-gradient-descent>
AUTHORS: Frederik Kunstner, Philipp Hennig, Lukas Balles
HIGHLIGHT: We dispute this argument by showing that the empirical Fisher---unlike the Fisher---does not generally capture second-order information.

375, TITLE: Flow-based Image-to-Image Translation with Feature Disentanglement
<https://papers.nips.cc/paper/8670-flow-based-image-to-image-translation-with-feature-disentanglement>
AUTHORS: Ruho Kondo, Keisuke Kawano, Satoshi Koide, Takuro Kutsuna
HIGHLIGHT: To this end we propose a flow-based image-to-image model, called Flow U-Net with Squeeze modules (FUNS), that allows us to disentangle the features while retaining the ability to generate highquality diverse images from condition images.

376, TITLE: Learning dynamic polynomial proofs
<https://papers.nips.cc/paper/8671-learning-dynamic-polynomial-proofs>
AUTHORS: Alhussein Fawzi, Mateusz Malinowski, Hamza Fawzi, Omar Fawzi
HIGHLIGHT: In this paper, we consider the fundamental computational task of automatically searching for proofs of polynomial inequalities.

377, TITLE: Shape and Time Distortion Loss for Training Deep Time Series Forecasting Models
<https://papers.nips.cc/paper/8672-shape-and-time-distortion-loss-for-training-deep-time-series-forecasting-models>
AUTHORS: Vincent LE GUEN, Nicolas THOME
HIGHLIGHT: To handle this challenging task, we introduce DILATE (DIstortion Loss including shApe and TimE), a new objective function for training deep neural networks.

378, TITLE: Understanding Attention and Generalization in Graph Neural Networks
<https://papers.nips.cc/paper/8673-understanding-attention-and-generalization-in-graph-neural-networks>
AUTHORS: Boris Knyazev, Graham W. Taylor, Mohamed Amer
HIGHLIGHT: Motivated by insights from the work on Graph Isomorphism Networks, we design simple graph reasoning tasks that allow us to study attention in a controlled environment.

379, TITLE: Data Cleansing for Models Trained with SGD
<https://papers.nips.cc/paper/8674-data-cleansing-for-models-trained-with-sgd>
AUTHORS: Satoshi Hara, Atsushi Nitanda, Takanori Maehara
HIGHLIGHT: In this paper, we propose an algorithm that can identify influential instances without using any domain knowledge.

380, TITLE: Curvilinear Distance Metric Learning
<https://papers.nips.cc/paper/8675-curvilinear-distance-metric-learning>
AUTHORS: Shuo Chen, Lei Luo, Jian Yang, Chen Gong, Jun Li, Heng Huang
HIGHLIGHT: After that, by extending such straight lines to general curved forms, we propose a Curvilinear Distance Metric Learning (CDML) method, which adaptively learns the nonlinear geometries of the training data.

381, TITLE: Embedding Symbolic Knowledge into Deep Networks
<https://papers.nips.cc/paper/8676-embedding-symbolic-knowledge-into-deep-networks>
AUTHORS: Xie Yaqi, Ziwei Xu, Kuldeep S Meel, Mohan Kankanhalli, Harold Soh
HIGHLIGHT: In this work, we aim to leverage prior symbolic knowledge to improve the performance of deep models.

382, TITLE: Modeling Uncertainty by Learning a Hierarchy of Deep Neural Connections
<https://papers.nips.cc/paper/8677-modeling-uncertainty-by-learning-a-hierarchy-of-deep-neural-connections>
AUTHORS: Raanan Yehezkel Rohekar, Yaniv Gurwicz, Shami Nisimov, Gal Novik
HIGHLIGHT: We propose an approach for modeling this confounder by sharing neural connectivity patterns between the generative and discriminative networks.

383, TITLE: Efficient Graph Generation with Graph Recurrent Attention Networks
<https://papers.nips.cc/paper/8678-efficient-graph-generation-with-graph-recurrent-attention-networks>
AUTHORS: Renjie Liao, Yujia Li, Yang Song, Shenlong Wang, Will Hamilton, David K. Duvenaud, Raquel Urtasun, Richard Zemel
HIGHLIGHT: We propose a new family of efficient and expressive deep generative models of graphs, called Graph Recurrent Attention Networks (GRANs).

384, TITLE: Beyond Alternating Updates for Matrix Factorization with Inertial Bregman Proximal Gradient Algorithms
<https://papers.nips.cc/paper/8679-beyond-alternating-updates-for-matrix-factorization-with-inertial-bregman-proximal-gradient-algorithms>
AUTHORS: Mahesh Chandra Mukkamala, Peter Ochs
HIGHLIGHT: We exploit this theory by proposing a novel Bregman distance for matrix factorization problems, which, at the same time, allows for simple/closed form update steps.

- 385, TITLE: Learning Deep Bilinear Transformation for Fine-grained Image Representation
<https://papers.nips.cc/paper/8680-learning-deep-bilinear-transformation-for-fine-grained-image-representation>
AUTHORS: Heliang Zheng, Jianlong Fu, Zheng-Jun Zha, Jiebo Luo
HIGHLIGHT: In this paper, we propose a deep bilinear transformation (DBT) block, which can be deeply stacked in convolutional neural networks to learn fine-grained image representations.
- 386, TITLE: Practical Deep Learning with Bayesian Principles
<https://papers.nips.cc/paper/8681-practical-deep-learning-with-bayesian-principles>
AUTHORS: Kazuki Osawa, Siddharth Swaroop, Mohammad Emtiyaz E. Khan, Anirudh Jain, Runa Eschenhagen, Richard E. Turner, Rio Yokota
HIGHLIGHT: In this paper, we demonstrate practical training of deep networks with natural-gradient variational inference.
- 387, TITLE: Training Language GANs from Scratch
<https://papers.nips.cc/paper/8682-training-language-gans-from-scratch>
AUTHORS: Cyprien de Masson d'Autume, Shakir Mohamed, Mihaela Rosca, Jack Rae
HIGHLIGHT: We combine existing techniques such as large batch sizes, dense rewards and discriminator regularization to stabilize and improve language GANs.
- 388, TITLE: Pseudo-Extended Markov chain Monte Carlo
<https://papers.nips.cc/paper/8683-pseudo-extended-markov-chain-monte-carlo>
AUTHORS: Christopher Nemeth, Fredrik Lindsten, Maurizio Filippone, James Hensman
HIGHLIGHT: In this paper, we introduce the pseudo-extended MCMC method as a simple approach for improving the mixing of the MCMC sampler for multi-modal posterior distributions.
- 389, TITLE: Differentially Private Bagging: Improved utility and cheaper privacy than subsample-and-aggregate
<https://papers.nips.cc/paper/8684-differentially-private-bagging-improved-utility-and-cheaper-privacy-than-subsample-and-aggregate>
AUTHORS: James Jordon, Jinsung Yoon, Mihaela van der Schaar
HIGHLIGHT: In this paper, we extend this approach by dividing the data several times (rather than just once) and learning models on each chunk within each division.
- 390, TITLE: Propagating Uncertainty in Reinforcement Learning via Wasserstein Barycenters
<https://papers.nips.cc/paper/8685-propagating-uncertainty-in-reinforcement-learning-via-wasserstein-barycenters>
AUTHORS: Alberto Maria Metelli, Amarildo Likmeta, Marcello Restelli
HIGHLIGHT: In this paper, we address this question by proposing a Bayesian framework in which we employ approximate posterior distributions to model the uncertainty of the value function and Wasserstein barycenters to propagate it across state-action pairs.
- 391, TITLE: On Adversarial Mixup Resynthesis
<https://papers.nips.cc/paper/8686-on-adversarial-mixup-resynthesis>
AUTHORS: Christopher Beckham, Sina Honari, Alex M. Lamb, Vikas Verma, Farnoosh Ghadiri, R Devon Hjelm, Yoshua Bengio, Chris Pal
HIGHLIGHT: In this paper, we explore new approaches to combining information encoded within the learned representations of auto-encoders.
- 392, TITLE: A Geometric Perspective on Optimal Representations for Reinforcement Learning
<https://papers.nips.cc/paper/8687-a-geometric-perspective-on-optimal-representations-for-reinforcement-learning>
AUTHORS: Marc Bellemare, Will Dabney, Robert Dadashi, Adrien Ali Taiga, Pablo Samuel Castro, Nicolas Le Roux, Dale Schuurmans, Tor Lattimore, Clare Lyle
HIGHLIGHT: We propose a new perspective on representation learning in reinforcement learning based on geometric properties of the space of value functions.
- 393, TITLE: Learning New Tricks From Old Dogs: Multi-Source Transfer Learning From Pre-Trained Networks
<https://papers.nips.cc/paper/8688-learning-new-tricks-from-old-dogs-multi-source-transfer-learning-from-pre-trained-networks>
AUTHORS: Joshua Lee, Prasanna Sattigeri, Gregory Wornell
HIGHLIGHT: For such scenarios, we consider the multi-source learning problem of training a classifier using an ensemble of pre-trained neural networks for a set of classes that have not been observed by any of the source networks, and for which we have very few training samples.
- 394, TITLE: Understanding and Improving Layer Normalization

- <https://papers.nips.cc/paper/8689-understanding-and-improving-layer-normalization>
AUTHORS: Jingjing Xu, Xu Sun, Zhiyuan Zhang, Guangxiang Zhao, Junyang Lin
HIGHLIGHT: In this paper, our main contribution is to take a step further in understanding LayerNorm.
- 395, TITLE: Uncertainty-based Continual Learning with Adaptive Regularization
<https://papers.nips.cc/paper/8690-uncertainty-based-continual-learning-with-adaptive-regularization>
AUTHORS: Hongjoon Ahn, Sungmin Cha, Donggyu Lee, Taesup Moon
HIGHLIGHT: We introduce a new neural network-based continual learning algorithm, dubbed as Uncertainty-regularized Continual Learning (UCL), which builds on traditional Bayesian online learning framework with variational inference.
- 396, TITLE: LIIR: Learning Individual Intrinsic Reward in Multi-Agent Reinforcement Learning
<https://papers.nips.cc/paper/8691-liir-learning-individual-intrinsic-reward-in-multi-agent-reinforcement-learning>
AUTHORS: Yali Du, Lei Han, Meng Fang, Ji Liu, Tianhong Dai, Dacheng Tao
HIGHLIGHT: In this paper, we propose to merge the two directions and learn each agent an intrinsic reward function which diversely stimulates the agents at each time step.
- 397, TITLE: U-Time: A Fully Convolutional Network for Time Series Segmentation Applied to Sleep Staging
<https://papers.nips.cc/paper/8692-u-time-a-fully-convolutional-network-for-time-series-segmentation-applied-to-sleep-staging>
AUTHORS: Mathias Perslev, Michael Jensen, Sune Darkner, Poul Jørgen Jennum, Christian Igel
HIGHLIGHT: We propose U-Time, a fully feed-forward deep learning approach to physiological time series segmentation developed for the analysis of sleep data.
- 398, TITLE: Massively scalable Sinkhorn distances via the Nystrom method
<https://papers.nips.cc/paper/8693-massively-scalable-sinkhorn-distances-via-the-nystrom-method>
AUTHORS: Jason Altschuler, Francis Bach, Alessandro Rudi, Jonathan Niles-Weed
HIGHLIGHT: In this work, we show that this challenge is surprisingly easy to circumvent: combining two simple techniques—the Nyström method and Sinkhorn scaling—provably yields an accurate approximation of the Sinkhorn distance with significantly lower time and memory requirements than other approaches.
- 399, TITLE: Double Quantization for Communication-Efficient Distributed Optimization
<https://papers.nips.cc/paper/8694-double-quantization-for-communication-efficient-distributed-optimization>
AUTHORS: Yue Yu, Jiaxiang Wu, Longbo Huang
HIGHLIGHT: In this paper, to reduce the communication complexity, we propose double quantization, a general scheme for quantizing both model parameters and gradients.
- 400, TITLE: Globally optimal score-based learning of directed acyclic graphs in high-dimensions
<https://papers.nips.cc/paper/8695-globally-optimal-score-based-learning-of-directed-acyclic-graphs-in-high-dimensions>
AUTHORS: Bryon Aragam, Arash Amini, Qing Zhou
HIGHLIGHT: We prove that $\Omega(s \log p)$ samples suffice to learn a sparse Gaussian directed acyclic graph (DAG) from data, where s is the maximum Markov blanket size.
- 401, TITLE: Multi-relational Poincare Graph Embeddings
<https://papers.nips.cc/paper/8696-multi-relational-poincare-graph-embeddings>
AUTHORS: Ivana Balazevic, Carl Allen, Timothy Hospedales
HIGHLIGHT: To address this, we propose a model that embeds multi-relational graph data in the Poincaré ball model of hyperbolic space.
- 402, TITLE: No-Press Diplomacy: Modeling Multi-Agent Gameplay
<https://papers.nips.cc/paper/8697-no-press-diplomacy-modeling-multi-agent-gameplay>
AUTHORS: Philip Paquette, Yuchen Lu, SETON STEVEN BOCCO, Max Smith, Satya O.-G., Jonathan K. Kummerfeld, Joelle Pineau, Satinder Singh, Aaron C. Courville
HIGHLIGHT: In this work, we focus on training an agent that learns to play the No Press version of Diplomacy where there is no dedicated communication channel between players.
- 403, TITLE: State Aggregation Learning from Markov Transition Data
<https://papers.nips.cc/paper/8698-state-aggregation-learning-from-markov-transition-data>
AUTHORS: Yaqi Duan, Tracy Ke, Mengdi Wang
HIGHLIGHT: In this paper, we propose a tractable algorithm that estimates the probabilistic aggregation map from the system's trajectory.

404, TITLE: Disentangling Influence: Using disentangled representations to audit model predictions
<https://papers.nips.cc/paper/8699-disentangling-influence-using-disentangled-representations-to-audit-model-predictions>
AUTHORS: Charles Marx, Richard Phillips, Sorelle Friedler, Carlos Scheidegger, Suresh Venkatasubramanian
HIGHLIGHT: In this paper, we develop disentangled influence audits, a procedure to audit the indirect influence of features.

405, TITLE: Successor Uncertainties: Exploration and Uncertainty in Temporal Difference Learning
<https://papers.nips.cc/paper/8700-successor-uncertainties-exploration-and-uncertainty-in-temporal-difference-learning>
AUTHORS: David Janz, Jiri Hron, Przemyslaw Mazur, Katja Hofmann, Jos? Miguel Hern?ndez-Lobato, Sebastian Tschatschek
HIGHLIGHT: We use these insights to design Successor Uncertainties (SU), a cheap and easy to implement RVF algorithm that retains key properties of PSRL.

406, TITLE: Partially Encrypted Deep Learning using Functional Encryption
<https://papers.nips.cc/paper/8701-partially-encrypted-deep-learning-using-functional-encryption>
AUTHORS: Th?o Ryffel, David Pointcheval, Francis Bach, Edouard Dufour-Sans, Romain Gay
HIGHLIGHT: We propose a practical framework to perform partially encrypted and privacy-preserving predictions which combines adversarial training and functional encryption.

407, TITLE: Decentralized Cooperative Stochastic Bandits
<https://papers.nips.cc/paper/8702-decentralized-cooperative-stochastic-bandits>
AUTHORS: David Mart?nez-Rubio, Varun Kanade, Patrick Rebeschini
HIGHLIGHT: We study a decentralized cooperative stochastic multi-armed bandit problem with K arms on a network of N agents.

408, TITLE: Statistical bounds for entropic optimal transport: sample complexity and the central limit theorem
<https://papers.nips.cc/paper/8703-statistical-bounds-for-entropic-optimal-transport-sample-complexity-and-the-central-limit-theorem>
AUTHORS: Gonzalo Mena, Jonathan Niles-Weed
HIGHLIGHT: We prove several fundamental statistical bounds for entropic OT with the squared Euclidean cost between subgaussian probability measures in arbitrary dimension.

409, TITLE: Efficient Deep Approximation of GMMs
<https://papers.nips.cc/paper/8704-efficient-deep-approximation-of-gmms>
AUTHORS: Shirin Jalali, Carl Nuzman, Iraj Saniee
HIGHLIGHT: In this work, we extend this idea to a rich class of functions, namely the discriminant functions that arise in optimal Bayesian classification of Gaussian mixture models (GMMs) in \mathbb{R}^n .

410, TITLE: Learning low-dimensional state embeddings and metastable clusters from time series data
<https://papers.nips.cc/paper/8705-learning-low-dimensional-state-embeddings-and-metastable-clusters-from-time-series-data>
AUTHORS: Yifan Sun, Yaqi Duan, Hao Gong, Mengdi Wang
HIGHLIGHT: In the spirit of diffusion map, we propose an efficient method for learning a low-dimensional state embedding and capturing the process's dynamics.

411, TITLE: Exploiting Local and Global Structure for Point Cloud Semantic Segmentation with Contextual Point Representations
<https://papers.nips.cc/paper/8706-exploiting-local-and-global-structure-for-point-cloud-semantic-segmentation-with-contextual-point-representations>
AUTHORS: Xu Wang, Jingming He, Lin Ma
HIGHLIGHT: In this paper, we propose one novel model for point cloud semantic segmentation, which exploits both the local and global structures within the point cloud based on the contextual point representations.

412, TITLE: Scalable Bayesian dynamic covariance modeling with variational Wishart and inverse Wishart processes
<https://papers.nips.cc/paper/8707-scalable-bayesian-dynamic-covariance-modeling-with-variational-wishart-and-inverse-wishart-processes>
AUTHORS: Creighton Heaukulani, Mark van der Wilk
HIGHLIGHT: We implement gradient-based variational inference routines for Wishart and inverse Wishart processes, which we apply as Bayesian models for the dynamic, heteroskedastic covariance matrix of a multivariate time series.

413, TITLE: Kernel Instrumental Variable Regression
<https://papers.nips.cc/paper/8708-kernel-instrumental-variable-regression>
AUTHORS: Rahul Singh, Maneesh Sahani, Arthur Gretton

HIGHLIGHT: We propose kernel instrumental variable regression (KIV), a nonparametric generalization of 2SLS, modeling relations among X , Y , and Z as nonlinear functions in reproducing kernel Hilbert spaces (RKHSs).

414, **TITLE:** Symmetry-Based Disentangled Representation Learning requires Interaction with Environments
<https://papers.nips.cc/paper/8709-symmetry-based-disentangled-representation-learning-requires-interaction-with-environments>
AUTHORS: Hugo Caselles-Dupr?, Michael Garcia Ortiz, David Filliat
HIGHLIGHT: We build on their work and make observations, theoretical and empirical, that lead us to argue that Symmetry-Based Disentangled Representation Learning cannot only be based on static observations: agents should interact with the environment to discover its symmetries.

415, **TITLE:** Fast Efficient Hyperparameter Tuning for Policy Gradient Methods
<https://papers.nips.cc/paper/8710-fast-efficient-hyperparameter-tuning-for-policy-gradient-methods>
AUTHORS: Supratik Paul, Vitaly Kurin, Shimon Whiteson
HIGHLIGHT: In this paper, we propose Hyperparameter Optimisation on the Fly (HOOF), a gradient-free algorithm that requires no more than one training run to automatically adapt the hyperparameter that affect the policy update directly through the gradient.

416, **TITLE:** Offline Contextual Bayesian Optimization
<https://papers.nips.cc/paper/8711-offline-contextual-bayesian-optimization>
AUTHORS: Ian Char, Youngseog Chung, Willie Neiswanger, Kirthevasan Kandasamy, Oak Nelson, Mark Boyer, Egemen Kolenen
HIGHLIGHT: In this work, we describe a theoretically grounded Bayesian optimization method to tackle this problem.

417, **TITLE:** Making the Cut: A Bandit-based Approach to Tiered Interviewing
<https://papers.nips.cc/paper/8712-making-the-cut-a-bandit-based-approach-to-tiered-interviewing>
AUTHORS: Candice Schumann, Zhi Lang, Jeffrey Foster, John Dickerson
HIGHLIGHT: We present new algorithms in both the probably approximately correct (PAC) and fixed-budget settings that select a near-optimal cohort with provable guarantees.

418, **TITLE:** Unsupervised Scalable Representation Learning for Multivariate Time Series
<https://papers.nips.cc/paper/8713-unsupervised-scalable-representation-learning-for-multivariate-time-series>
AUTHORS: Jean-Yves Franceschi, Aymeric Dieuleveut, Martin Jaggi
HIGHLIGHT: In this paper, we tackle this challenge by proposing an unsupervised method to learn universal embeddings of time series.

419, **TITLE:** A state-space model for inferring effective connectivity of latent neural dynamics from simultaneous EEG/fMRI
<https://papers.nips.cc/paper/8714-a-state-space-model-for-inferring-effective-connectivity-of-latent-neural-dynamics-from-simultaneous-eegefMRI>
AUTHORS: Tao Tu, John Paisley, Stefan Haufe, Paul Sajda
HIGHLIGHT: In this study, we develop a linear state-space model to infer the effective connectivity in a distributed brain network based on simultaneously recorded EEG and fMRI data.

420, **TITLE:** End to end learning and optimization on graphs
<https://papers.nips.cc/paper/8715-end-to-end-learning-and-optimization-on-graphs>
AUTHORS: Bryan Wilder, Eric Ewing, Bistra Dilkina, Milind Tambe
HIGHLIGHT: Here, we propose an alternative decision-focused learning approach that integrates a differentiable proxy for common graph optimization problems as a layer in learned systems.

421, **TITLE:** Game Design for Eliciting Distinguishable Behavior
<https://papers.nips.cc/paper/8716-game-design-for-eliciting-distinguishable-behavior>
AUTHORS: Fan Yang, Liu Leqi, Yifan Wu, Zachary Lipton, Pradeep K. Ravikumar, Tom M. Mitchell, William W. Cohen
HIGHLIGHT: In this paper, we formulate the task of designing behavior diagnostic games that elicit distinguishable behavior as a mutual information maximization problem, which can be solved by optimizing a variational lower bound.

422, **TITLE:** When does label smoothing help?
<https://papers.nips.cc/paper/8717-when-does-label-smoothing-help>
AUTHORS: Rafael M?ller, Simon Kornblith, Geoffrey E. Hinton
HIGHLIGHT: We show that label smoothing encourages the representations of training examples from the same class to group in tight clusters.

- 423, TITLE: Finite-Time Performance Bounds and Adaptive Learning Rate Selection for Two Time-Scale Reinforcement Learning
<https://papers.nips.cc/paper/8718-finite-time-performance-bounds-and-adaptive-learning-rate-selection-for-two-time-scale-reinforcement-learning>
AUTHORS: Harsh Gupta, R. Srikant, Lei Ying
HIGHLIGHT: We present finite-time performance bounds for the case where the learning rate is fixed.
- 424, TITLE: Rethinking Deep Neural Network Ownership Verification: Embedding Passports to Defeat Ambiguity Attacks
<https://papers.nips.cc/paper/8719-rethinking-deep-neural-network-ownership-verification-embedding-passports-to-defeat-ambiguity-attacks>
AUTHORS: Lixin Fan, Kam Woh Ng, Chee Seng Chan
HIGHLIGHT: As remedies to the above-mentioned loophole, this paper proposes novel passport-based DNN ownership verification schemes which are both robust to network modifications and resilient to ambiguity attacks.
- 425, TITLE: Scalable Spike Source Localization in Extracellular Recordings using Amortized Variational Inference
<https://papers.nips.cc/paper/8720-scalable-spike-source-localization-in-extracellular-recordings-using-amortized-variational-inference>
AUTHORS: Cole Hurwitz, Kai Xu, Akash Srivastava, Alessio Buccino, Matthias Hennig
HIGHLIGHT: In this work, we present a Bayesian modelling approach for localizing the source of individual spikes on high-density, microelectrode arrays.
- 426, TITLE: Optimal Sketching for Kronecker Product Regression and Low Rank Approximation
<https://papers.nips.cc/paper/8721-optimal-sketching-for-kronecker-product-regression-and-low-rank-approximation>
AUTHORS: Huaian Diao, Rajesh Jayaram, Zhao Song, Wen Sun, David Woodruff
HIGHLIGHT: In this work, we provide significantly faster algorithms.
- 427, TITLE: Distribution-Independent PAC Learning of Halfspaces with Massart Noise
<https://papers.nips.cc/paper/8722-distribution-independent-pac-learning-of-halfspaces-with-massart-noise>
AUTHORS: Ilias Diakonikolas, Themis Gouleakis, Christos Tzamos
HIGHLIGHT: We study the problem of distribution-independent PAC learning of halfspaces in the presence of Massart noise.
- 428, TITLE: The Convergence Rate of Neural Networks for Learned Functions of Different Frequencies
<https://papers.nips.cc/paper/8723-the-convergence-rate-of-neural-networks-for-learned-functions-of-different-frequencies>
AUTHORS: Basri Ronen, David Jacobs, Yoni Kasten, Shira Kritchman
HIGHLIGHT: We study the relationship between the frequency of a function and the speed at which a neural network learns it.
- 429, TITLE: Adaptive Auxiliary Task Weighting for Reinforcement Learning
<https://papers.nips.cc/paper/8724-adaptive-auxiliary-task-weighting-for-reinforcement-learning>
AUTHORS: Xingyu Lin, Harjatin Baweja, George Kantor, David Held
HIGHLIGHT: In this work, we propose a principled online learning algorithm that dynamically combines different auxiliary tasks to speed up training for reinforcement learning.
- 430, TITLE: Blocking Bandits
<https://papers.nips.cc/paper/8725-blocking-bandits>
AUTHORS: Soumya Basu, Rajat Sen, Sujay Sanghavi, Sanjay Shakkottai
HIGHLIGHT: We consider a novel stochastic multi-armed bandit setting, where playing an arm makes it unavailable for a fixed number of time slots thereafter.
- 431, TITLE: Global Convergence of Least Squares EM for Demixing Two Log-Concave Densities
<https://papers.nips.cc/paper/8726-global-convergence-of-least-squares-em-for-demixing-two-log-concave-densities>
AUTHORS: Wei Qian, Yuqian Zhang, Yudong Chen
HIGHLIGHT: We demonstrate that Least Squares EM, a variant of the EM algorithm, converges to the true location parameter from a randomly initialized point.
- 432, TITLE: Prior-Free Dynamic Auctions with Low Regret Buyers
<https://papers.nips.cc/paper/8727-prior-free-dynamic-auctions-with-low-regret-buyers>
AUTHORS: Yuan Deng, Jon Schneider, Balasubramanian Sivan
HIGHLIGHT: In this work, we do away with this assumption and consider the prior-free setting where the buyer's value each round is chosen adversarially (possibly adaptively).

- 433, TITLE: On Single Source Robustness in Deep Fusion Models
<https://papers.nips.cc/paper/8728-on-single-source-robustness-in-deep-fusion-models>
AUTHORS: Taewan Kim, Joydeep Ghosh
HIGHLIGHT: Motivated by this discovery, two possible approaches are proposed to increase robustness: a carefully designed loss with corresponding training algorithms for deep fusion models, and a simple convolutional fusion layer that has a structural advantage in dealing with noise.
- 434, TITLE: Policy Evaluation with Latent Confounders via Optimal Balance
<https://papers.nips.cc/paper/8729-policy-evaluation-with-latent-confounders-via-optimal-balance>
AUTHORS: Andrew Bennett, Nathan Kallus
HIGHLIGHT: Instead, we propose an adversarial objective and weights that minimize it, ensuring sufficient balance in the latent confounders regardless of outcome model.
- 435, TITLE: Think Globally, Act Locally: A Deep Neural Network Approach to High-Dimensional Time Series Forecasting
<https://papers.nips.cc/paper/8730-think-globally-act-locally-a-deep-neural-network-approach-to-high-dimensional-time-series-forecasting>
AUTHORS: Rajat Sen, Hsiang-Fu Yu, Inderjit S. Dhillon
HIGHLIGHT: In this paper, we seek to correct this deficiency and propose DeepGLO, a deep forecasting model which thinks globally and acts locally.
- 436, TITLE: Adaptive Cross-Modal Few-shot Learning
<https://papers.nips.cc/paper/8731-adaptive-cross-modal-few-shot-learning>
AUTHORS: Chen Xing, Negar Rostamzadeh, Boris Oreshkin, Pedro O. O. Pinheiro
HIGHLIGHT: In this paper, we propose to leverage cross-modal information to enhance metric-based few-shot learning methods.
- 437, TITLE: Spectral Modification of Graphs for Improved Spectral Clustering
<https://papers.nips.cc/paper/8732-spectral-modification-of-graphs-for-improved-spectral-clustering>
AUTHORS: Ioannis Koutis, Huong Le
HIGHLIGHT: In this paper we show that for any graph G , there exists a 'spectral maximizer' graph H which is cut-similar to G , but has eigenvalues that are near the theoretical limit implied by the cut structure of G .
- 438, TITLE: Hyperbolic Graph Convolutional Neural Networks
<https://papers.nips.cc/paper/8733-hyperbolic-graph-convolutional-neural-networks>
AUTHORS: Ines Chami, Zhitao Ying, Christopher R?, Jure Leskovec
HIGHLIGHT: Here we propose Hyperbolic Graph Convolutional Neural Network (HGNC), the first inductive hyperbolic GCN that leverages both the expressiveness of GCNs and hyperbolic geometry to learn inductive node representations for hierarchical and scale-free graphs.
- 439, TITLE: Cost Effective Active Search
<https://papers.nips.cc/paper/8734-cost-effective-active-search>
AUTHORS: Shali Jiang, Roman Garnett, Benjamin Moseley
HIGHLIGHT: We propose simple and fast approximations for computing its expectation, which serves as an essential role in our proposed policy.
- 440, TITLE: Exploration Bonus for Regret Minimization in Discrete and Continuous Average Reward MDPs
<https://papers.nips.cc/paper/8735-exploration-bonus-for-regret-minimization-in-discrete-and-continuous-average-reward-mdps>
AUTHORS: Jian QIAN, Ronan Fruit, Matteo Pirota, Alessandro Lazaric
HIGHLIGHT: While it has been analyzed in infinite-horizon discounted and finite-horizon problems, we focus on designing and analysing the exploration bonus in the more challenging infinite-horizon undiscounted setting.
- 441, TITLE: Hybrid 8-bit Floating Point (HFP8) Training and Inference for Deep Neural Networks
<https://papers.nips.cc/paper/8736-hybrid-8-bit-floating-point-hfp8-training-and-inference-for-deep-neural-networks>
AUTHORS: Xiao Sun, Jungwook Choi, Chia-Yu Chen, Naigang Wang, Swagath Venkataramani, Vijayalakshmi (Viji) Srinivasan, Xiaodong Cui, Wei Zhang, Kailash Gopalakrishnan
HIGHLIGHT: Using theoretical insights, we propose a hybrid FP8 (HFP8) format and DNN end-to-end distributed training procedure.
- 442, TITLE: Tight Certificates of Adversarial Robustness for Randomly Smoothed Classifiers
<https://papers.nips.cc/paper/8737-tight-certificates-of-adversarial-robustness-for-randomly-smoothed-classifiers>
AUTHORS: Guang-He Lee, Yang Yuan, Shiyu Chang, Tommi Jaakkola

HIGHLIGHT: In particular, we offer adversarial robustness guarantees and associated algorithms for the discrete case where the adversary is ℓ_0 bounded.

443, TITLE: Poisson-Minibatching for Gibbs Sampling with Convergence Rate Guarantees
<https://papers.nips.cc/paper/8738-poisson-minibatching-for-gibbs-sampling-with-convergence-rate-guarantees>

AUTHORS: Ruqi Zhang, Christopher M. De Sa

HIGHLIGHT: In this paper, we propose a new auxiliary-variable minibatched Gibbs sampling method, Poisson-minibatching Gibbs, which both produces unbiased samples and has a theoretical guarantee on its convergence rate.

444, TITLE: One ticket to win them all: generalizing lottery ticket initializations across datasets and optimizers
<https://papers.nips.cc/paper/8739-one-ticket-to-win-them-all-generalizing-lottery-ticket-initializations-across-datasets-and-optimizers>

AUTHORS: Ari Morcos, Haonan Yu, Michela Paganini, Yuandong Tian

HIGHLIGHT: Here, we attempt to answer this question by generating winning tickets for one training configuration (optimizer and dataset) and evaluating their performance on another configuration.

445, TITLE: Breaking the Glass Ceiling for Embedding-Based Classifiers for Large Output Spaces
<https://papers.nips.cc/paper/8740-breaking-the-glass-ceiling-for-embedding-based-classifiers-for-large-output-spaces>

AUTHORS: Chuan Guo, Ali Mousavi, Xiang Wu, Daniel N. Holtmann-Rice, Satyen Kale, Sashank Reddi, Sanjiv Kumar

HIGHLIGHT: In this paper, we demonstrate that theoretically there is no limitation to using low-dimensional embedding-based methods, and provide experimental evidence that overfitting is the root cause of the poor performance of embedding-based methods.

446, TITLE: Fair Algorithms for Clustering
<https://papers.nips.cc/paper/8741-fair-algorithms-for-clustering>

AUTHORS: Suman Bera, Deeparnab Chakrabarty, Nicolas Flores, Maryam Negahbani

HIGHLIGHT: We study the problem of finding low-cost ℓ_{∞} fair clusterings in data where each data point may belong to many protected groups.

447, TITLE: Learning Mean-Field Games
<https://papers.nips.cc/paper/8742-learning-mean-field-games>

AUTHORS: Xin Guo, Anran Hu, Renyuan Xu, Junzi Zhang

HIGHLIGHT: This paper presents a general mean-field game (GMFG) framework for simultaneous learning and decision-making in stochastic games with a large population.

448, TITLE: SpArSe: Sparse Architecture Search for CNNs on Resource-Constrained Microcontrollers
<https://papers.nips.cc/paper/8743-sparse-sparse-architecture-search-for-cnns-on-resource-constrained-microcontrollers>

AUTHORS: Igor Fedorov, Ryan P. Adams, Matthew Mattina, Paul Whatmough

HIGHLIGHT: This paper challenges the idea that CNNs are not suitable for deployment on MCUs. We demonstrate that it is possible to automatically design CNNs which generalize well, while also being small enough to fit onto memory-limited MCUs.

449, TITLE: Deep imitation learning for molecular inverse problems
<https://papers.nips.cc/paper/8744-deep-imitation-learning-for-molecular-inverse-problems>

AUTHORS: Eric Jonas

HIGHLIGHT: We treat this as a problem of graph-structured prediction, where armed with per-vertex information on a subset of the vertices, we infer the edges and edge types.

450, TITLE: Visual Concept-Metaconcept Learning
<https://papers.nips.cc/paper/8745-visual-concept-metaconcept-learning>

AUTHORS: Chi Han, Jiayuan Mao, Chuang Gan, Josh Tenenbaum, Jiajun Wu

HIGHLIGHT: In this paper, we propose the visual concept-metaconcept learner (VCML) for joint learning of concepts and metaconcepts from images and associated question-answer pairs.

451, TITLE: Few-shot Video-to-Video Synthesis
<https://papers.nips.cc/paper/8746-few-shot-video-to-video-synthesis>

AUTHORS: Ting-Chun Wang, Ming-Yu Liu, Andrew Tao, Guilin Liu, Bryan Catanzaro, Jan Kautz

HIGHLIGHT: To address the limitations, we propose a few-shot vid2vid framework, which learns to synthesize videos of previously unseen subjects or scenes by leveraging few example images of the target at test time.

452, TITLE: Neural Similarity Learning
<https://papers.nips.cc/paper/8747-neural-similarity-learning>

- AUTHORS: Weiyang Liu, Zhen Liu, James M. Rehg, Le Song
HIGHLIGHT: By generalizing inner product with a bilinear matrix, we propose the neural similarity which serves as a learnable parametric similarity measure for CNNs.
- 453, TITLE: Ordered Memory
<https://papers.nips.cc/paper/8748-ordered-memory>
AUTHORS: Yikang Shen, Shawn Tan, Arian Hosseini, Zhouhan Lin, Alessandro Sordoni, Aaron C. Courville
HIGHLIGHT: In this paper, we propose the Ordered Memory architecture.
- 454, TITLE: MixMatch: A Holistic Approach to Semi-Supervised Learning
<https://papers.nips.cc/paper/8749-mixmatch-a-holistic-approach-to-semi-supervised-learning>
AUTHORS: David Berthelot, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, Colin A. Raffel
HIGHLIGHT: In this work, we unify the current dominant approaches for semi-supervised learning to produce a new algorithm, MixMatch, that guesses low-entropy labels for data-augmented unlabeled examples and mixes labeled and unlabeled data using MixUp.
- 455, TITLE: Multivariate Triangular Quantile Maps for Novelty Detection
<https://papers.nips.cc/paper/8750-multivariate-triangular-quantile-maps-for-novelty-detection>
AUTHORS: Jingjing Wang, Sun Sun, Yaoliang Yu
HIGHLIGHT: In this work, we present a general framework for neural novelty detection that centers around a multivariate extension of the univariate quantile function.
- 456, TITLE: Fast Parallel Algorithms for Statistical Subset Selection Problems
<https://papers.nips.cc/paper/8751-fast-parallel-algorithms-for-statistical-subset-selection-problems>
AUTHORS: Sharon Qian, Yaron Singer
HIGHLIGHT: In this paper, we propose a new framework for designing fast parallel algorithms for fundamental statistical subset selection tasks that include feature selection and experimental design.
- 457, TITLE: PHYRE: A New Benchmark for Physical Reasoning
<https://papers.nips.cc/paper/8752-phyre-a-new-benchmark-for-physical-reasoning>
AUTHORS: Anton Bakhtin, Laurens van der Maaten, Justin Johnson, Laura Gustafson, Ross Girshick
HIGHLIGHT: We develop the PHYRE benchmark for physical reasoning that contains a set of simple classical mechanics puzzles in a 2D physical environment.
- 458, TITLE: On the number of variables to use in principal component regression
<https://papers.nips.cc/paper/8753-on-the-number-of-variables-to-use-in-principal-component-regression>
AUTHORS: Ji Xu, Daniel J. Hsu
HIGHLIGHT: We study least squares linear regression over uncorrelated Gaussian features that are selected in order of decreasing variance.
- 459, TITLE: Factor Group-Sparse Regularization for Efficient Low-Rank Matrix Recovery
<https://papers.nips.cc/paper/8754-factor-group-sparse-regularization-for-efficient-low-rank-matrix-recovery>
AUTHORS: Jicong Fan, Lijun Ding, Yudong Chen, Madeleine Udell
HIGHLIGHT: This paper develops a new class of nonconvex regularizers for low-rank matrix recovery.
- 460, TITLE: Mutually Regressive Point Processes
<https://papers.nips.cc/paper/8755-mutually-regressive-point-processes>
AUTHORS: Ifigenia Apostolopoulou, Scott Linderman, Kyle Miller, Artur Dubrawski
HIGHLIGHT: In this paper, we introduce the first general class of Bayesian point process models extended with a nonlinear component that allows both excitatory and inhibitory relationships in continuous time.
- 461, TITLE: Data-driven Estimation of Sinusoid Frequencies
<https://papers.nips.cc/paper/8756-data-driven-estimation-of-sinusoid-frequencies>
AUTHORS: Gautier Izacard, Sreyas Mohan, Carlos Fernandez-Granda
HIGHLIGHT: In this work, we propose a novel neural-network architecture that produces a significantly more accurate representation, and combine it with an additional neural-network module trained to detect the number of frequencies.
- 462, TITLE: E2-Train: Training State-of-the-art CNNs with Over 80% Less Energy
<https://papers.nips.cc/paper/8757-e2-train-training-state-of-the-art-cnns-with-over-80-less-energy>
AUTHORS: Ziyu Jiang, Yue Wang, Xiaohan Chen, Pengfei Xu, Yang Zhao, Yingyan Lin, Zhangyang Wang

HIGHLIGHT: This paper attempts to explore an orthogonal direction: how to conduct more energy-efficient training of CNNs, so as to enable on-device training?

463, **TITLE:** ANODEV2: A Coupled Neural ODE Framework
<https://papers.nips.cc/paper/8758-anodev2-a-coupled-neural-ode-framework>
AUTHORS: Tianjun Zhang, Zhewei Yao, Amir Gholami, Joseph E. Gonzalez, Kurt Keutzer, Michael W. Mahoney, George Biros
HIGHLIGHT: Here, we propose OURS, which extends this approach by introducing a framework that allows ODE-based evolution for both the weights and the activations, in a coupled formulation.

464, **TITLE:** Estimating Entropy of Distributions in Constant Space
<https://papers.nips.cc/paper/8759-estimating-entropy-of-distributions-in-constant-space>
AUTHORS: Jayadev Acharya, Sourbh Bhadane, Piotr Indyk, Ziteng Sun
HIGHLIGHT: Our main contribution is an algorithm that requires $O\left(\frac{k \log(1/\epsilon)^2}{\epsilon^3}\right)$ samples and a constant $O(1)$ memory words of space and outputs a $\pm \epsilon$ estimate of $H(p)$.

465, **TITLE:** On the Utility of Learning about Humans for Human-AI Coordination
<https://papers.nips.cc/paper/8760-on-the-utility-of-learning-about-humans-for-human-ai-coordination>
AUTHORS: Micah Carroll, Rohin Shah, Mark K. Ho, Tom Griffiths, Sanjit Seshia, Pieter Abbeel, Anca Dragan
HIGHLIGHT: To demonstrate this, we introduce a simple environment that requires challenging coordination, based on the popular game Overcooked, and learn a simple model that mimics human play.

466, **TITLE:** Efficient Regret Minimization Algorithm for Extensive-Form Correlated Equilibrium
<https://papers.nips.cc/paper/8761-efficient-regret-minimization-algorithm-for-extensive-form-correlated-equilibrium>
AUTHORS: Gabriele Farina, Chun Kai Ling, Fei Fang, Tuomas Sandholm
HIGHLIGHT: In this paper, we introduce the first efficient regret minimization algorithm for computing extensive-form correlated equilibria in large two-player general-sum games with no chance moves.

467, **TITLE:** Learning in Generalized Linear Contextual Bandits with Stochastic Delays
<https://papers.nips.cc/paper/8762-learning-in-generalized-linear-contextual-bandits-with-stochastic-delays>
AUTHORS: Zhengyuan Zhou, Renyuan Xu, Jose Blanchet
HIGHLIGHT: In this paper, we consider online learning in generalized linear contextual bandits where rewards are not immediately observed.

468, **TITLE:** Empirically Measuring Concentration: Fundamental Limits on Intrinsic Robustness
<https://papers.nips.cc/paper/8763-empirically-measuring-concentration-fundamental-limits-on-intrinsic-robustness>
AUTHORS: Saeed Mahloujifar, Xiao Zhang, Mohammad Mahmoody, David Evans
HIGHLIGHT: This paper presents a method for empirically measuring and bounding the concentration of a concrete dataset which is proven to converge to the actual concentration.

469, **TITLE:** Optimistic Regret Minimization for Extensive-Form Games via Dilated Distance-Generating Functions
<https://papers.nips.cc/paper/8764-optimistic-regret-minimization-for-extensive-form-games-via-dilated-distance-generating-functions>
AUTHORS: Gabriele Farina, Christian Kroer, Tuomas Sandholm
HIGHLIGHT: We study the performance of optimistic regret-minimization algorithms for both minimizing regret in, and computing Nash equilibria of, zero-sum extensive-form games.

470, **TITLE:** Learning Non-Convergent Non-Persistent Short-Run MCMC Toward Energy-Based Model
<https://papers.nips.cc/paper/8765-learning-non-convergent-non-persistent-short-run-mcmc-toward-energy-based-model>
AUTHORS: Erik Nijkamp, Mitch Hill, Song-Chun Zhu, Ying Nian Wu
HIGHLIGHT: This paper studies a curious phenomenon in learning energy-based model (EBM) using MCMC.

471, **TITLE:** Enhancing the Locality and Breaking the Memory Bottleneck of Transformer on Time Series Forecasting
<https://papers.nips.cc/paper/8766-enhancing-the-locality-and-breaking-the-memory-bottleneck-of-transformer-on-time-series-forecasting>
AUTHORS: Shiyang Li, Xiaoyong Jin, Yao Xuan, Xiyu Zhou, Wenhui Chen, Yu-Xiang Wang, Xifeng Yan
HIGHLIGHT: In this paper, we propose to tackle such forecasting problem with Transformer.

472, **TITLE:** On the Accuracy of Influence Functions for Measuring Group Effects
<https://papers.nips.cc/paper/8767-on-the-accuracy-of-influence-functions-for-measuring-group-effects>

- AUTHORS: Pang Wei W. Koh, Kai-Siang Ang, Hubert Teo, Percy S. Liang
HIGHLIGHT: In this paper, we find that across many different types of groups and for a range of real-world datasets, the predicted effect (using influence functions) of a group correlates surprisingly well with its actual effect, even if the absolute and relative errors are large.
- 473, TITLE: Face Reconstruction from Voice using Generative Adversarial Networks
<https://papers.nips.cc/paper/8768-face-reconstruction-from-voice-using-generative-adversarial-networks>
AUTHORS: Yandong Wen, Bhiksha Raj, Rita Singh
HIGHLIGHT: In this paper, we address the challenge posed by a subtask of voice profiling - reconstructing someone's face from their voice.
- 474, TITLE: Incremental Few-Shot Learning with Attention Attractor Networks
<https://papers.nips.cc/paper/8769-incremental-few-shot-learning-with-attention-attractor-networks>
AUTHORS: Mengye Ren, Renjie Liao, Ethan Fetaya, Richard Zemel
HIGHLIGHT: To this end, we propose a meta-learning model, the Attention Attractor Network, which regularizes the learning of novel classes.
- 475, TITLE: On Testing for Biases in Peer Review
<https://papers.nips.cc/paper/8770-on-testing-for-biases-in-peer-review>
AUTHORS: Ivan Stelmakh, Nihar Shah, Aarti Singh
HIGHLIGHT: We consider the issue of biases in scholarly research, specifically, in peer review.
- 476, TITLE: Learning Disentangled Representation for Robust Person Re-identification
<https://papers.nips.cc/paper/8771-learning-disentangled-representation-for-robust-person-re-identification>
AUTHORS: Chanho Eom, Bumsub Ham
HIGHLIGHT: To tackle this problem, we propose to disentangle identity-related and -unrelated features from person images.
- 477, TITLE: Balancing Efficiency and Fairness in On-Demand Ridesourcing
<https://papers.nips.cc/paper/8772-balancing-efficiency-and-fairness-in-on-demand-ridesourcing>
AUTHORS: Nixie S. Lesmana, Xuan Zhang, Xiaohui Bei
HIGHLIGHT: In this paper, we focus on both the system efficiency and the fairness among drivers and quantitatively analyze the trade-offs between these two objectives.
- 478, TITLE: Latent Ordinary Differential Equations for Irregularly-Sampled Time Series
<https://papers.nips.cc/paper/8773-latent-ordinary-differential-equations-for-irregularly-sampled-time-series>
AUTHORS: Yulia Rubanova, Tian Qi Chen, David K. Duvenaud
HIGHLIGHT: We generalize RNNs to have continuous-time hidden dynamics defined by ordinary differential equations (ODEs), a model we call ODE-RNNs.
- 479, TITLE: Deep RGB-D Canonical Correlation Analysis For Sparse Depth Completion
<https://papers.nips.cc/paper/8774-deep-rgb-d-canonical-correlation-analysis-for-sparse-depth-completion>
AUTHORS: Yiqi Zhong, Cho-Ying Wu, Suyu You, Ulrich Neumann
HIGHLIGHT: In this paper, we propose our Correlation For Completion Network (CFCNet), an end-to-end deep learning model that uses the correlation between two data sources to perform sparse depth completion.
- 480, TITLE: Input Similarity from the Neural Network Perspective
<https://papers.nips.cc/paper/8775-input-similarity-from-the-neural-network-perspective>
AUTHORS: Guillaume Charpiat, Nicolas Girard, Loris Felardos, Yuliya Tarabalka
HIGHLIGHT: Given a trained neural network, we aim at understanding how similar it considers any two samples.
- 481, TITLE: Adaptive Sequence Submodularity
<https://papers.nips.cc/paper/8776-adaptive-sequence-submodularity>
AUTHORS: Marko Mitrovic, Ehsan Kazemi, Moran Feldman, Andreas Krause, Amin Karbasi
HIGHLIGHT: In this paper, we view the problem of adaptive and sequential decision making through the lens of submodularity and propose an adaptive greedy policy with strong theoretical guarantees.
- 482, TITLE: Weight Agnostic Neural Networks
<https://papers.nips.cc/paper/8777-weight-agnostic-neural-networks>
AUTHORS: Adam Gaier, David Ha

HIGHLIGHT: In this work, we question to what extent neural network architectures alone, without learning any weight parameters, can encode solutions for a given task.

483, **TITLE:** Learning to Predict Without Looking Ahead: World Models Without Forward Prediction
<https://papers.nips.cc/paper/8778-learning-to-predict-without-looking-ahead-world-models-without-forward-prediction>

AUTHORS: Daniel Freeman, David Ha, Luke Metz

HIGHLIGHT: In this work, we introduce a modification to traditional reinforcement learning which we call observational dropout, whereby we limit the agents ability to observe the real environment at each timestep.

484, **TITLE:** Reducing the variance in online optimization by transporting past gradients

<https://papers.nips.cc/paper/8779-reducing-the-variance-in-online-optimization-by-transporting-past-gradients>

AUTHORS: S'bastien Arnold, Pierre-Antoine Manzagol, Reza Babanezhad Harikandeh, Ioannis Mitliagkas, Nicolas Le Roux

HIGHLIGHT: We propose to correct this staleness using the idea of $\{\em implicit gradient transport\}$ (IGT) which transforms gradients computed at previous iterates into gradients evaluated at the current iterate without using the Hessian explicitly.

485, **TITLE:** Characterizing Bias in Classifiers using Generative Models

<https://papers.nips.cc/paper/8780-characterizing-bias-in-classifiers-using-generative-models>

AUTHORS: Daniel McDuff, Shuang Ma, Yale Song, Ashish Kapoor

HIGHLIGHT: We propose a simulation-based approach for interrogating classifiers using generative adversarial models in a systematic manner.

486, **TITLE:** Optimal Stochastic and Online Learning with Individual Iterates

<https://papers.nips.cc/paper/8781-optimal-stochastic-and-online-learning-with-individual-iterates>

AUTHORS: Yunwen Lei, Peng Yang, Ke Tang, Ding-Xuan Zhou

HIGHLIGHT: In this paper, we propose a theoretically sound strategy to select an individual iterate of the vanilla SCMD, which is able to achieve optimal rates for both convex and strongly convex problems in a non-smooth learning setting.

487, **TITLE:** Policy Learning for Fairness in Ranking

<https://papers.nips.cc/paper/8782-policy-learning-for-fairness-in-ranking>

AUTHORS: Ashudeep Singh, Thorsten Joachims

HIGHLIGHT: To address this need, we propose a general LTR framework that can optimize a wide range of utility metrics (e.g. NDCG) while satisfying fairness of exposure constraints with respect to the items.

488, **TITLE:** Off-Policy Evaluation via Off-Policy Classification

<https://papers.nips.cc/paper/8783-off-policy-evaluation-via-off-policy-classification>

AUTHORS: Alexander Irpan, Kanishka Rao, Konstantinos Bousmalis, Chris Harris, Julian Ibarz, Sergey Levine

HIGHLIGHT: In this work, we consider the problem of model selection for deep reinforcement learning (RL) in real-world environments.

489, **TITLE:** Regularized Gradient Boosting

<https://papers.nips.cc/paper/8784-regularized-gradient-boosting>

AUTHORS: Corinna Cortes, Mehryar Mohri, Dmitry Storcheus

HIGHLIGHT: We introduce a new algorithm, called rgb , that directly benefits from these generalization bounds and that, at every boosting round, applies the $\{\em Structural Risk Minimization\}$ principle to search for a base predictor with the best empirical fit versus complexity trade-off.

490, **TITLE:** Efficient Probabilistic Inference in the Quest for Physics Beyond the Standard Model

<https://papers.nips.cc/paper/8785-efficient-probabilistic-inference-in-the-quest-for-physics-beyond-the-standard-model>

AUTHORS: Atilim Gunes Baydin, Lei Shao, Wahid Bhimji, Lukas Heinrich, Saeid Naderiparizi, Andreas Munk, Jialin Liu, Bradley Gram-Hansen, Gilles Louppe, Lawrence Meadows, Philip Torr, Victor Lee, Kyle Cranmer, Mr. Prabhat, Frank Wood

HIGHLIGHT: We present a novel probabilistic programming framework that couples directly to existing large-scale simulators through a cross-platform probabilistic execution protocol, which allows general-purpose inference engines to record and control random number draws within simulators in a language-agnostic way.

491, **TITLE:** Markov Random Fields for Collaborative Filtering

<https://papers.nips.cc/paper/8786-markov-random-fields-for-collaborative-filtering>

AUTHORS: Harald Steck

HIGHLIGHT: In this paper, we model the dependencies among the items that are recommended to a user in a collaborative-filtering problem via a Gaussian Markov Random Field (MRF).

492, TITLE: A Step Toward Quantifying Independently Reproducible Machine Learning Research
<https://papers.nips.cc/paper/8787-a-step-toward-quantifying-independently-reproducible-machine-learning-research>
AUTHORS: Edward Raff
HIGHLIGHT: We take the first step toward a quantifiable answer by manually attempting to implement 255 papers published from 1984 until 2017, recording features of each paper, and performing statistical analysis of the results.

493, TITLE: Scalable Global Optimization via Local Bayesian Optimization
<https://papers.nips.cc/paper/8788-scalable-global-optimization-via-local-bayesian-optimization>
AUTHORS: David Eriksson, Michael Pearce, Jacob Gardner, Ryan D. Turner, Matthias Poloczek
HIGHLIGHT: We propose the TuRBO algorithm that fits a collection of local models and performs a principled global allocation of samples across these models via an implicit bandit approach.

494, TITLE: Time-series Generative Adversarial Networks
<https://papers.nips.cc/paper/8789-time-series-generative-adversarial-networks>
AUTHORS: Jinsung Yoon, Daniel Jarrett, M Van Der Schaar
HIGHLIGHT: We propose a novel framework for generating realistic time-series data that combines the flexibility of the unsupervised paradigm with the control afforded by supervised training.

495, TITLE: Ouroboros: On Accelerating Training of Transformer-Based Language Models
<https://papers.nips.cc/paper/8790-ouroboros-on-accelerating-training-of-transformer-based-language-models>
AUTHORS: Qian Yang, Zhouyuan Huo, Wenlin Wang, Lawrence Carin
HIGHLIGHT: We propose the first model-parallel algorithm that speeds the training of Transformer-based language models.

496, TITLE: A Refined Margin Distribution Analysis for Forest Representation Learning
<https://papers.nips.cc/paper/8791-a-refined-margin-distribution-analysis-for-forest-representation-learning>
AUTHORS: Shen-Huan Lyu, Liang Yang, Zhi-Hua Zhou
HIGHLIGHT: In this paper, we formulate the forest representation learning approach called CasDF as an additive model which boosts the augmented feature instead of the prediction.

497, TITLE: Robustness to Adversarial Perturbations in Learning from Incomplete Data
<https://papers.nips.cc/paper/8792-robustness-to-adversarial-perturbations-in-learning-from-incomplete-data>
AUTHORS: Amir Najafi, Shin-ichi Maeda, Masanori Koyama, Takeru Miyato
HIGHLIGHT: We develop a generalization theory for our framework based on a number of novel complexity measures, such as an adversarial extension of Rademacher complexity and its semi-supervised analogue.

498, TITLE: Exploring Unexplored Tensor Network Decompositions for Convolutional Neural Networks
<https://papers.nips.cc/paper/8793-exploring-unexplored-tensor-network-decompositions-for-convolutional-neural-networks>
AUTHORS: Kohei Hayashi, Taiki Yamaguchi, Yohei Sugawara, Shin-ichi Maeda
HIGHLIGHT: In this study, we first characterize a decomposition class specific to CNNs by adopting a flexible graphical notation.

499, TITLE: An Adaptive Empirical Bayesian Method for Sparse Deep Learning
<https://papers.nips.cc/paper/8794-an-adaptive-empirical-bayesian-method-for-sparse-deep-learning>
AUTHORS: Wei Deng, Xiao Zhang, Faming Liang, Guang Lin
HIGHLIGHT: We propose a novel adaptive empirical Bayesian (AEB) method for sparse deep learning, where the sparsity is ensured via a class of self-adaptive spike-and-slab priors.

500, TITLE: Adaptive Influence Maximization with Myopic Feedback
<https://papers.nips.cc/paper/8795-adaptive-influence-maximization-with-myopic-feedback>
AUTHORS: Binghui Peng, Wei Chen
HIGHLIGHT: We study the adaptive influence maximization problem with myopic feedback under the independent cascade model: one sequentially selects k nodes as seeds one by one from a social network, and each selected seed returns the immediate neighbors it activates as the feedback available for by later selections, and the goal is to maximize the expected number of total activated nodes, referred as the influence spread.

501, TITLE: Focused Quantization for Sparse CNNs
<https://papers.nips.cc/paper/8796-focused-quantization-for-sparse-cnns>
AUTHORS: Yiren Zhao, Xitong Gao, Daniel Bates, Robert Mullins, Cheng-Zhong Xu
HIGHLIGHT: In this paper, we attend to the statistical properties of sparse CNNs and present focused quantization, a novel quantization strategy based on power-of-two values, which exploits the weight distributions after fine-grained pruning.

502, TITLE: Quantum Embedding of Knowledge for Reasoning
<https://papers.nips.cc/paper/8797-quantum-embedding-of-knowledge-for-reasoning>
AUTHORS: Dinesh Garg, Shajith Ikkal Mohamed, Santosh K. Srivastava, Harit Vishwakarma, Hima Karanam, L Venkata Subramaniam
HIGHLIGHT: We present a novel approach called Embed2Reason (E2R) that embeds a symbolic KB into a vector space in a logical structure preserving manner.

503, TITLE: Optimal Best Markovian Arm Identification with Fixed Confidence
<https://papers.nips.cc/paper/8798-optimal-best-markovian-arm-identification-with-fixed-confidence>
AUTHORS: Vrettos Moulos
HIGHLIGHT: We give a complete characterization of the sampling complexity of best Markovian arm identification in one-parameter Markovian bandit models.

504, TITLE: Limiting Extrapolation in Linear Approximate Value Iteration
<https://papers.nips.cc/paper/8799-limiting-extrapolation-in-linear-approximate-value-iteration>
AUTHORS: Andrea Zanette, Alessandro Lazaric, Mykel J. Kochenderfer, Emma Brunskill
HIGHLIGHT: We introduce an algorithm that approximates value functions by combining Q-values estimated at a set of anchor states.

505, TITLE: Almost Horizon-Free Structure-Aware Best Policy Identification with a Generative Model
<https://papers.nips.cc/paper/8800-almost-horizon-free-structure-aware-best-policy-identification-with-a-generative-model>
AUTHORS: Andrea Zanette, Mykel J. Kochenderfer, Emma Brunskill
HIGHLIGHT: We propose an algorithm that is initially agnostic to the MDP but that can leverage the specific MDP structure, expressed in terms of variances of the rewards and next-state value function, and gaps in the optimal action-value function to reduce the sample complexity needed to find a good policy, precisely highlighting the contribution of each state-action pair to the final sample complexity.

506, TITLE: Invertible Convolutional Flow
<https://papers.nips.cc/paper/8801-invertible-convolutional-flow>
AUTHORS: Mahdi Karami, Dale Schuurmans, Jascha Sohl-Dickstein, Laurent Dinh, Daniel Duckworth
HIGHLIGHT: As an alternative, we investigate a set of novel normalizing flows based on the circular and symmetric convolutions.

507, TITLE: A Latent Variational Framework for Stochastic Optimization
<https://papers.nips.cc/paper/8802-a-latent-variational-framework-for-stochastic-optimization>
AUTHORS: Philippe Casgrain
HIGHLIGHT: This paper provides a unifying theoretical framework for stochastic optimization algorithms by means of a latent stochastic variational problem.

508, TITLE: Topology-Preserving Deep Image Segmentation
<https://papers.nips.cc/paper/8803-topology-preserving-deep-image-segmentation>
AUTHORS: Xiaoling Hu, Fuxin Li, Dimitris Samaras, Chao Chen
HIGHLIGHT: We propose a novel method that learns to segment with correct topology.

509, TITLE: Connective Cognition Network for Directional Visual Commonsense Reasoning
<https://papers.nips.cc/paper/8804-connective-cognition-network-for-directional-visual-commonsense-reasoning>
AUTHORS: Aming Wu, Linchao Zhu, Yahong Han, Yi Yang
HIGHLIGHT: Inspired by this idea, towards VCR, we propose a connective cognition network (CCN) to dynamically reorganize the visual neuron connectivity that is contextualized by the meaning of questions and answers.

510, TITLE: Online Markov Decoding: Lower Bounds and Near-Optimal Approximation Algorithms
<https://papers.nips.cc/paper/8805-online-markov-decoding-lower-bounds-and-near-optimal-approximation-algorithms>
AUTHORS: Vikas Garg, Tamar Pichkhadze
HIGHLIGHT: We resolve the fundamental problem of online decoding with general n th order ergodic Markov chain models.

511, TITLE: A Meta-MDP Approach to Exploration for Lifelong Reinforcement Learning
<https://papers.nips.cc/paper/8806-a-meta-mdp-approach-to-exploration-for-lifelong-reinforcement-learning>
AUTHORS: Francisco Garcia, Philip S. Thomas

HIGHLIGHT: In this paper we consider the problem of how a reinforcement learning agent that is tasked with solving a sequence of reinforcement learning problems (a sequence of Markov decision processes) can use knowledge acquired early in its lifetime to improve its ability to solve new problems.

512, **TITLE:** Push-pull Feedback Implements Hierarchical Information Retrieval Efficiently
<https://papers.nips.cc/paper/8807-push-pull-feedback-implements-hierarchical-information-retrieval-efficiently>
AUTHORS: Xiao Liu, Xiaolong Zou, Zilong Ji, Gengshuo Tian, Yuanyuan Mi, Tiejun Huang, K. Y. Michael Wong, Si Wu
HIGHLIGHT: Here, we investigate the role of feedback in hierarchical information retrieval.

513, **TITLE:** Learning Disentangled Representations for Recommendation
<https://papers.nips.cc/paper/8808-learning-disentangled-representations-for-recommendation>
AUTHORS: Jianxin Ma, Chang Zhou, Peng Cui, Hongxia Yang, Wenwu Zhu
HIGHLIGHT: In this paper, we present the MACRo-mlcro Disentangled Variational Auto-Encoder (MacridVAE) for learning disentangled representations from user behavior.

514, **TITLE:** Graph Neural Tangent Kernel: Fusing Graph Neural Networks with Graph Kernels
<https://papers.nips.cc/paper/8809-graph-neural-tangent-kernel-fusing-graph-neural-networks-with-graph-kernels>
AUTHORS: Simon S. Du, Kangcheng Hou, Russ R. Salakhutdinov, Barnabas Poczos, Ruosong Wang, Keyulu Xu
HIGHLIGHT: The current paper presents a new class of graph kernels, Graph Neural Tangent Kernels (GNTKs), which correspond to infinitely wide multi-layer GNNs trained by gradient descent.

515, **TITLE:** In-Place Zero-Space Memory Protection for CNN
<https://papers.nips.cc/paper/8810-in-place-zero-space-memory-protection-for-cnn>
AUTHORS: Hui Guan, Lin Ning, Zhen Lin, Xipeng Shen, Huiyang Zhou, Seung-Hwan Lim
HIGHLIGHT: This paper introduces in-place zero-space ECC assisted with a new training scheme weight distribution-oriented training.

516, **TITLE:** Acceleration via Symplectic Discretization of High-Resolution Differential Equations
<https://papers.nips.cc/paper/8811-acceleration-via-symplectic-discretization-of-high-resolution-differential-equations>
AUTHORS: Bin Shi, Simon S. Du, Weijie Su, Michael I. Jordan
HIGHLIGHT: We study first-order optimization algorithms obtained by discretizing ordinary differential equations (ODEs) corresponding to Nesterov's accelerated gradient methods (NAGs) and Polyak's heavy-ball method.

517, **TITLE:** XLNet: Generalized Autoregressive Pretraining for Language Understanding
<https://papers.nips.cc/paper/8812-xl-net-generalized-autoregressive-pretraining-for-language-understanding>
AUTHORS: Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R. Salakhutdinov, Quoc V. Le
HIGHLIGHT: In light of these pros and cons, we propose XLNet, a generalized autoregressive pretraining method that (1) enables learning bidirectional contexts by maximizing the expected likelihood over all permutations of the factorization order and (2) overcomes the limitations of BERT thanks to its autoregressive formulation.

518, **TITLE:** Comparison Against Task Driven Artificial Neural Networks Reveals Functional Properties in Mouse Visual Cortex
<https://papers.nips.cc/paper/8813-comparison-against-task-driven-artificial-neural-networks-reveals-functional-properties-in-mouse-visual-cortex>
AUTHORS: Jianghong Shi, Eric Shea-Brown, Michael Buice
HIGHLIGHT: We find that the comparison procedure is robust to different choices of stimuli set and the level of sub-sampling that one might expect in a large scale brain survey with thousands of neurons.

519, **TITLE:** Variance Reduced Policy Evaluation with Smooth Function Approximation
<https://papers.nips.cc/paper/8814-variance-reduced-policy-evaluation-with-smooth-function-approximation>
AUTHORS: Hoi-To Wai, Mingyi Hong, Zhuoran Yang, Zhaoran Wang, Kexin Tang
HIGHLIGHT: We formulate the policy evaluation problem as a non-convex primal-dual, finite-sum optimization problem, whose primal sub-problem is non-convex and dual sub-problem is strongly concave.

520, **TITLE:** Learning GANs and Ensembles Using Discrepancy
<https://papers.nips.cc/paper/8815-learning-gans-and-ensembles-using-discrepancy>
AUTHORS: Ben Adlam, Corinna Cortes, Mehryar Mohri, Ningshan Zhang
HIGHLIGHT: We present efficient algorithms using discrepancy for two tasks: training a GAN directly, namely DGAN, and mixing previously trained generative models, namely EDGAN.

- 521, TITLE: Co-Generation with GANs using AIS based HMC
<https://papers.nips.cc/paper/8816-co-generation-with-gans-using-ais-based-hmc>
AUTHORS: Tiantian Fang, Alexander Schwing
HIGHLIGHT: Therefore, in this paper, we study the occurring challenges for co-generation with GANs.
- 522, TITLE: AttentionXML: Label Tree-based Attention-Aware Deep Model for High-Performance Extreme Multi-Label Text Classification
<https://papers.nips.cc/paper/8817-attentionxml-label-tree-based-attention-aware-deep-model-for-high-performance-extreme-multi-label-text-classification>
AUTHORS: Ronghui You, Zihan Zhang, Ziye Wang, Suyang Dai, Hiroshi Mamitsuka, Shanfeng Zhu
HIGHLIGHT: We propose a new label tree-based deep learning model for XMTC, called AttentionXML, with two unique features: 1) a multi-label attention mechanism with raw text as input, which allows to capture the most relevant part of text to each label; and 2) a shallow and wide probabilistic label tree (PLT), which allows to handle millions of labels, especially for "tail labels".
- 523, TITLE: Addressing Sample Complexity in Visual Tasks Using HER and Hallucinatory GANs
<https://papers.nips.cc/paper/8818-addressing-sample-complexity-in-visual-tasks-using-her-and-hallucinatory-gans>
AUTHORS: Himanshu Sahni, Toby Buckley, Pieter Abbeel, Ilya Kuzovkin
HIGHLIGHT: In this work, we show how visual trajectories can be hallucinated to appear successful by altering agent observations using a generative model trained on relatively few snapshots of the goal.
- 524, TITLE: Abstract Reasoning with Distracting Features
<https://papers.nips.cc/paper/8819-abstract-reasoning-with-distracting-features>
AUTHORS: Kecheng Zheng, Zheng-Jun Zha, Wei Wei
HIGHLIGHT: Inspired this fact, we propose feature robust abstract reasoning (FRAR) model, which consists of a reinforcement learning based teacher network to determine the sequence of training and a student network for predictions.
- 525, TITLE: Generalized Block-Diagonal Structure Pursuit: Learning Soft Latent Task Assignment against Negative Transfer
<https://papers.nips.cc/paper/8820-generalized-block-diagonal-structure-pursuit-learning-soft-latent-task-assignment-against-negative-transfer>
AUTHORS: Zhiyong Yang, Qianqian Xu, Yangbangyan Jiang, Xiaochun Cao, Qingming Huang
HIGHLIGHT: To circumvent this issue, we propose a novel multi-task learning method, which simultaneously learns latent task representations and a block-diagonal Latent Task Assignment Matrix (LTAM).
- 526, TITLE: Adversarial Training and Robustness for Multiple Perturbations
<https://papers.nips.cc/paper/8821-adversarial-training-and-robustness-for-multiple-perturbations>
AUTHORS: Florian Tramer, Dan Boneh
HIGHLIGHT: Our aim is to understand the reasons underlying this robustness trade-off, and to train models that are simultaneously robust to multiple perturbation types.
- 527, TITLE: Doubly-Robust Lasso Bandit
<https://papers.nips.cc/paper/8822-doubly-robust-lasso-bandit>
AUTHORS: Gi-Soo Kim, Myunghee Cho Paik
HIGHLIGHT: We consider the stochastic linear contextual bandit problem and propose a novel algorithm, namely the Doubly-Robust Lasso Bandit algorithm, which exploits the sparse structure of the regression parameter as in Lasso, while blending the doubly-robust technique used in missing data literature.
- 528, TITLE: DM2C: Deep Mixed-Modal Clustering
<https://papers.nips.cc/paper/8823-dm2c-deep-mixed-modal-clustering>
AUTHORS: Yangbangyan Jiang, Qianqian Xu, Zhiyong Yang, Xiaochun Cao, Qingming Huang
HIGHLIGHT: In this paper, we consider a more challenging task where each instance is represented in only one modality, which we call mixed-modal data.
- 529, TITLE: MaCow: Masked Convolutional Generative Flow
<https://papers.nips.cc/paper/8824-macow-masked-convolutional-generative-flow>
AUTHORS: Xuezhe Ma, Xiang Kong, Shanghang Zhang, Eduard Hovy
HIGHLIGHT: In this work, we introduce masked convolutional generative flow (MaCow), a simple yet effective architecture of generative flow using masked convolution.
- 530, TITLE: Learning by Abstraction: The Neural State Machine
<https://papers.nips.cc/paper/8825-learning-by-abstraction-the-neural-state-machine>

AUTHORS: Drew Hudson, Christopher D. Manning
HIGHLIGHT: We introduce the Neural State Machine, seeking to bridge the gap between the neural and symbolic views of AI and integrate their complementary strengths for the task of visual reasoning.

531, TITLE: Adaptive Gradient-Based Meta-Learning Methods
<https://papers.nips.cc/paper/8826-adaptive-gradient-based-meta-learning-methods>
AUTHORS: Mikhail Khodak, Maria-Florina F. Balcan, Ameet S. Talwalkar
HIGHLIGHT: We build a theoretical framework for designing and understanding practical meta-learning methods that integrates sophisticated formalizations of task-similarity with the extensive literature on online convex optimization and sequential prediction algorithms.

532, TITLE: Equipping Experts/Bandits with Long-term Memory
<https://papers.nips.cc/paper/8827-equipping-expertsbandits-with-long-term-memory>
AUTHORS: Kai Zheng, Haipeng Luo, Ilias Diakonikolas, Liwei Wang
HIGHLIGHT: We propose the first black-box approach to obtaining long-term memory guarantees for online learning in the sense of Bousquet and Warmuth, 2002, by reducing the problem to achieving typical switching regret.

533, TITLE: A Regularized Approach to Sparse Optimal Policy in Reinforcement Learning
<https://papers.nips.cc/paper/8828-a-regularized-approach-to-sparse-optimal-policy-in-reinforcement-learning>
AUTHORS: Wenhao Yang, Xiang Li, Zhihua Zhang
HIGHLIGHT: We propose and study a general framework for regularized Markov decision processes (MDPs) where the goal is to find an optimal policy that maximizes the expected discounted total reward plus a policy regularization term.

534, TITLE: Scalable inference of topic evolution via models for latent geometric structures
<https://papers.nips.cc/paper/8829-scalable-inference-of-topic-evolution-via-models-for-latent-geometric-structures>
AUTHORS: Mikhail Yurochkin, Zhiwei Fan, Aritra Guha, Paraschos Koutris, XuanLong Nguyen
HIGHLIGHT: We develop new models and algorithms for learning the temporal dynamics of the topic polytopes and related geometric objects that arise in topic model based inference.

535, TITLE: Effective End-to-end Unsupervised Outlier Detection via Inlier Priority of Discriminative Network
<https://papers.nips.cc/paper/8830-effective-end-to-end-unsupervised-outlier-detection-via-inlier-priority-of-discriminative-network>
AUTHORS: Siqi Wang, Yijie Zeng, Xinwang Liu, En Zhu, Jianping Yin, Chuanfu Xu, Marius Kloft
HIGHLIGHT: In this paper, we propose a framework named E³Outlier, which can perform UOD in a both effective and end-to-end manner: First, instead of the commonly-used autoencoders in previous end-to-end UOD methods, E³Outlier for the first time leverages a discriminative DNN for better representation learning, by using surrogate supervision to create multiple pseudo classes from original unlabelled data.

536, TITLE: Deep Active Learning with a Neural Architecture Search
<https://papers.nips.cc/paper/8831-deep-active-learning-with-a-neural-architecture-search>
AUTHORS: Yonatan Geifman, Ran El-Yaniv
HIGHLIGHT: We challenge this assumption and propose a novel active strategy whereby the learning algorithm searches for effective architectures on the fly, while actively learning.

537, TITLE: Efficiently escaping saddle points on manifolds
<https://papers.nips.cc/paper/8832-efficiently-escaping-saddle-points-on-manifolds>
AUTHORS: Christopher Criscitiello, Nicolas Boumal
HIGHLIGHT: Generalizing Jin et al.'s recent work on perturbed gradient descent (PGD) for optimization on linear spaces [How to Escape Saddle Points Efficiently (2017), Stochastic Gradient Descent Escapes Saddle Points Efficiently (2019)], we study a version of perturbed Riemannian gradient descent (PRGD) to show that necessary optimality conditions can be met approximately with high probability, without evaluating the Hessian.

538, TITLE: AutoAssist: A Framework to Accelerate Training of Deep Neural Networks
<https://papers.nips.cc/paper/8833-autoassist-a-framework-to-accelerate-training-of-deep-neural-networks>
AUTHORS: Jiong Zhang, Hsiang-Fu Yu, Inderjit S. Dhillon
HIGHLIGHT: In this paper, we propose AutoAssist, a simple framework to accelerate training of a deep neural network.

539, TITLE: DFNet: Spectral CNNs for Graphs with Feedback-Looped Filters
<https://papers.nips.cc/paper/8834-dfnet-spectral-cnns-for-graphs-with-feedback-looped-filters>
AUTHORS: W. O. K. Asiri Suranga Wijesinghe, Qing Wang
HIGHLIGHT: We propose a novel spectral convolutional neural network (CNN) model on graph structured data, namely Distributed Feedback-Looped Networks (DFNets).

- 540, TITLE: Learning Dynamics of Attention: Human Prior for Interpretable Machine Reasoning
<https://papers.nips.cc/paper/8835-learning-dynamics-of-attention-human-prior-for-interpretable-machine-reasoning>
AUTHORS: Wonjae Kim, Yoonho Lee
HIGHLIGHT: We propose Dynamics of Attention for Focus Transition (DAFT) as a human prior for machine reasoning.
- 541, TITLE: Comparing Unsupervised Word Translation Methods Step by Step
<https://papers.nips.cc/paper/8836-comparing-unsupervised-word-translation-methods-step-by-step>
AUTHORS: Mareike Hartmann, Yova Kementchedjieva, Anders Søgaard
HIGHLIGHT: We focus on the first step and compare distribution matching techniques in the context of language pairs for which mixed training stability and evaluation scores have been reported.
- 542, TITLE: Learning from Bad Data via Generation
<https://papers.nips.cc/paper/8837-learning-from-bad-data-via-generation>
AUTHORS: Tianyu Guo, Chang Xu, Boxin Shi, Chao Xu, Dacheng Tao
HIGHLIGHT: We suppose the real data distribution lies in a distribution set supported by the empirical distribution of bad data. A worst-case formulation can be developed over this distribution set, and then be interpreted as a generation task in an adversarial manner.
- 543, TITLE: Constrained deep neural network architecture search for IoT devices accounting for hardware calibration
<https://papers.nips.cc/paper/8838-constrained-deep-neural-network-architecture-search-for-iot-devices-accounting-for-hardware-calibration>
AUTHORS: Florian Scheidegger, Luca Benini, Costas Bekas, A. Cristiano I. Malossi
HIGHLIGHT: We propose a unique narrow-space architecture search that focuses on delivering low-cost and rapidly executing networks that respect strict memory and time requirements typical of Internet-of-Things (IoT) near-sensor computing platforms.
- 544, TITLE: Quantum Entropy Scoring for Fast Robust Mean Estimation and Improved Outlier Detection
<https://papers.nips.cc/paper/8839-quantum-entropy-scoring-for-fast-robust-mean-estimation-and-improved-outlier-detection>
AUTHORS: Yihe Dong, Samuel Hopkins, Jerry Li
HIGHLIGHT: We study two problems in high-dimensional robust statistics: $\text{\emph{robust mean estimation}}$ and $\text{\emph{outlier detection}}$.
- 545, TITLE: Iterative Least Trimmed Squares for Mixed Linear Regression
<https://papers.nips.cc/paper/8840-iterative-least-trimmed-squares-for-mixed-linear-regression>
AUTHORS: Yanyao Shen, Sujay Sanghavi
HIGHLIGHT: In this paper, we analyze ILTS in the setting of mixed linear regression with corruptions (MLR-C).
- 546, TITLE: Dynamic Ensemble Modeling Approach to Nonstationary Neural Decoding in Brain-Computer Interfaces
<https://papers.nips.cc/paper/8841-dynamic-ensemble-modeling-approach-to-nonstationary-neural-decoding-in-brain-computer-interfaces>
AUTHORS: Yu Qi, Bin Liu, Yueming Wang, Gang Pan
HIGHLIGHT: We propose a dynamic ensemble modeling (DyEnsemble) approach that is capable of adapting to changes in neural signals by employing a proper combination of decoding functions.
- 547, TITLE: Divergence-Augmented Policy Optimization
<https://papers.nips.cc/paper/8842-divergence-augmented-policy-optimization>
AUTHORS: Qing Wang, Yingru Li, Jiechao Xiong, Tong Zhang
HIGHLIGHT: This paper introduces a method to stabilize policy optimization when off-policy data are reused.
- 548, TITLE: Intrinsic dimension of data representations in deep neural networks
<https://papers.nips.cc/paper/8843-intrinsic-dimension-of-data-representations-in-deep-neural-networks>
AUTHORS: Alessio Ansuini, Alessandro Laio, Jakob H. Macke, Davide Zoccolan
HIGHLIGHT: Here we study the intrinsic dimensionality (ID) of data representations, i.e. the minimal number of parameters needed to describe a representation.
- 549, TITLE: Towards a Zero-One Law for Column Subset Selection
<https://papers.nips.cc/paper/8844-towards-a-zero-one-law-for-column-subset-selection>
AUTHORS: Zhao Song, David Woodruff, Peilin Zhong

HIGHLIGHT: In this work we give approximation algorithms for f function g which is approximately monotone and satisfies an approximate triangle inequality, and we show both of these conditions are necessary.

550, TITLE: Compositional De-Attention Networks
<https://papers.nips.cc/paper/8845-compositional-de-attention-networks>
AUTHORS: Yi Tay, Anh Tuan Luu, Aston Zhang, Shuohang Wang, Siu Cheung Hui
HIGHLIGHT: This paper proposes a new quasi-attention that is compositional in nature, i.e., learning whether to add , subtract or nullify a certain vector when learning representations.

551, TITLE: Dual Adversarial Semantics-Consistent Network for Generalized Zero-Shot Learning
<https://papers.nips.cc/paper/8846-dual-adversarial-semantics-consistent-network-for-generalized-zero-shot-learning>
AUTHORS: Jian Ni, Shanghang Zhang, Haiyong Xie
HIGHLIGHT: To address these limitations, we propose a Dual Adversarial Semantics-Consistent Network (referred to as DASCN), which learns both primal and dual Generative Adversarial Networks (GANs) in a unified framework for GZSL.

552, TITLE: Learning and Generalization in Overparameterized Neural Networks, Going Beyond Two Layers
<https://papers.nips.cc/paper/8847-learning-and-generalization-in-overparameterized-neural-networks-going-beyond-two-layers>
AUTHORS: Zeyuan Allen-Zhu, Yuanzhi Li, Yingyu Liang
HIGHLIGHT: In this work, we prove that overparameterized neural networks can learn some notable concept classes, including two and three-layer networks with fewer parameters and smooth activations.

553, TITLE: Mining GOLD Samples for Conditional GANs
<https://papers.nips.cc/paper/8848-mining-gold-samples-for-conditional-gans>
AUTHORS: Sangwoo Mo, Chiheon Kim, Sungwoong Kim, Minsu Cho, Jinwoo Shin
HIGHLIGHT: We introduce a simple yet effective approach to improving cGANs by measuring the discrepancy between the data distribution and the model distribution on given samples.

554, TITLE: Deep Model Transferability from Attribution Maps
<https://papers.nips.cc/paper/8849-deep-model-transferability-from-attribution-maps>
AUTHORS: Jie Song, Yixin Chen, Xinchao Wang, Chengchao Shen, Mingli Song
HIGHLIGHT: In this paper, we propose an embarrassingly simple yet very efficacious approach to estimating the transferability of deep networks, especially those handling vision tasks.

555, TITLE: Fully Parameterized Quantile Function for Distributional Reinforcement Learning
<https://papers.nips.cc/paper/8850-fully-parameterized-quantile-function-for-distributional-reinforcement-learning>
AUTHORS: Derek Yang, Li Zhao, Zichuan Lin, Tao Qin, Jiang Bian, Tie-Yan Liu
HIGHLIGHT: In this paper, we propose fully parameterized quantile function that parameterizes both the quantile fraction axis (i.e., the x-axis) and the value axis (i.e., y-axis) for distributional RL.

556, TITLE: Direct Optimization through $\arg \max$ for Discrete Variational Auto-Encoder
<https://papers.nips.cc/paper/8851-direct-optimization-through-arg-max-for-discrete-variational-auto-encoder>
AUTHORS: Guy Lorbom, Tommi Jaakkola, Andreea Gane, Tamir Hazan
HIGHLIGHT: In contrast to previous works which resort to softmax -based relaxations, we propose to optimize it directly by applying the $\text{direct loss minimization}$ approach.

557, TITLE: Distributional Reward Decomposition for Reinforcement Learning
<https://papers.nips.cc/paper/8852-distributional-reward-decomposition-for-reinforcement-learning>
AUTHORS: Zichuan Lin, Li Zhao, Derek Yang, Tao Qin, Tie-Yan Liu, Guangwen Yang
HIGHLIGHT: In this paper, we propose Distributional Reward Decomposition for Reinforcement Learning (DRDRL), a novel reward decomposition algorithm which captures the multiple reward channel structure under distributional setting.

558, TITLE: L_{DMI} : A Novel Information-theoretic Loss Function for Training Deep Nets Robust to Label Noise
https://papers.nips.cc/paper/8853-l_dmi-a-novel-information-theoretic-loss-function-for-training-deep-nets-robust-to-label-noise
AUTHORS: Yilun Xu, Peng Cao, Yuqing Kong, Yizhou Wang
HIGHLIGHT: In this paper, we propose a novel information-theoretic loss function, L_{DMI} , for training deep neural networks robust to label noise.

559, TITLE: Convergence Guarantees for Adaptive Bayesian Quadrature Methods
<https://papers.nips.cc/paper/8854-convergence-guarantees-for-adaptive-bayesian-quadrature-methods>
AUTHORS: Motonobu Kanagawa, Philipp Hennig

HIGHLIGHT: In this work, for a broad class of adaptive Bayesian quadrature methods, we prove consistency, deriving non-tight but informative convergence rates.

560, **TITLE:** Progressive Augmentation of GANs
<https://papers.nips.cc/paper/8855-progressive-augmentation-of-gans>
AUTHORS: Dan Zhang, Anna Khoreva
HIGHLIGHT: To mitigate this issue we introduce a new regularization technique - progressive augmentation of GANs (PA-GAN).

561, **TITLE:** UniXGrad: A Universal, Adaptive Algorithm with Optimal Guarantees for Constrained Optimization
<https://papers.nips.cc/paper/8856-unixgrad-a-universal-adaptive-algorithm-with-optimal-guarantees-for-constrained-optimization>
AUTHORS: Ali Kavis, Kfir Y. Levy, Francis Bach, Volkan Cevher
HIGHLIGHT: We propose a novel adaptive, accelerated algorithm for the stochastic constrained convex optimization setting. Our method, which is inspired by the Mirror-Prox method, achieves the optimal rates for smooth/non-smooth problems with either deterministic/stochastic first-order oracles.

562, **TITLE:** Meta-Surrogate Benchmarking for Hyperparameter Optimization
<https://papers.nips.cc/paper/8857-meta-surrogate-benchmarking-for-hyperparameter-optimization>
AUTHORS: Aaron Klein, Zhenwen Dai, Frank Hutter, Neil Lawrence, Javier Gonzalez
HIGHLIGHT: This work proposes a method to alleviate these issues by means of a meta-surrogate model for HPO tasks trained on off-line generated data.

563, **TITLE:** Learning to Perform Local Rewriting for Combinatorial Optimization
<https://papers.nips.cc/paper/8858-learning-to-perform-local-rewriting-for-combinatorial-optimization>
AUTHORS: Xinyun Chen, Yuandong Tian
HIGHLIGHT: In this paper, we propose NeuRewriter that learns a policy to pick heuristics and rewrite the local components of the current solution to iteratively improve it until convergence.

564, **TITLE:** Anti-efficient encoding in emergent communication
<https://papers.nips.cc/paper/8859-anti-efficient-encoding-in-emergent-communication>
AUTHORS: Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, Marco Baroni
HIGHLIGHT: We study whether the same pattern emerges when two neural networks, a "speaker" and a "listener", are trained to play a signaling game.

565, **TITLE:** Singleshot : a scalable Tucker tensor decomposition
<https://papers.nips.cc/paper/8860-singleshot-a-scalable-tucker-tensor-decomposition>
AUTHORS: Abraham Traore, Maxime Berar, Alain Rakotomamonjy
HIGHLIGHT: This paper introduces a new approach for the scalable Tucker decomposition problem.

566, **TITLE:** Neural Machine Translation with Soft Prototype
<https://papers.nips.cc/paper/8861-neural-machine-translation-with-soft-prototype>
AUTHORS: Yiren Wang, Yingce Xia, Fei Tian, Fei Gao, Tao Qin, Cheng Xiang Zhai, Tie-Yan Liu
HIGHLIGHT: In this work, we propose a new framework that introduces a soft prototype into the encoder-decoder architecture, which allows the decoder to have indirect access to both past and future information, such that each target word can be generated based on the better global understanding.

567, **TITLE:** Reliable training and estimation of variance networks
<https://papers.nips.cc/paper/8862-reliable-training-and-estimation-of-variance-networks>
AUTHORS: Nicki Skafte, Martin Jørgensen, Søren Hauberg
HIGHLIGHT: We propose and investigate new complementary methodologies for estimating predictive variance networks in regression neural networks.

568, **TITLE:** Copula Multi-label Learning
<https://papers.nips.cc/paper/8863-copula-multi-label-learning>
AUTHORS: Weiwei Liu
HIGHLIGHT: In particular, the paper first leverages the kernel trick to construct continuous distribution in the output space, and then estimates our proposed model semiparametrically where the copula is modeled parametrically, while the marginal distributions are modeled nonparametrically.

569, **TITLE:** Bayesian Learning of Sum-Product Networks

<https://papers.nips.cc/paper/8864-bayesian-learning-of-sum-product-networks>

AUTHORS: Martin Trapp, Robert Peharz, Hong Ge, Franz Pernkopf, Zoubin Ghahramani
HIGHLIGHT: In this paper, we introduce a well-principled Bayesian framework for SPN structure learning.

570, TITLE: Bayesian Batch Active Learning as Sparse Subset Approximation

<https://papers.nips.cc/paper/8865-bayesian-batch-active-learning-as-sparse-subset-approximation>

AUTHORS: Robert Pinsler, Jonathan Gordon, Eric Nalisnick, Jos? Miguel Hern?ndez-Lobato
HIGHLIGHT: In this paper, we introduce a novel Bayesian batch active learning approach that mitigates these issues.

571, TITLE: Optimal Sparsity-Sensitive Bounds for Distributed Mean Estimation

<https://papers.nips.cc/paper/8866-optimal-sparsity-sensitive-bounds-for-distributed-mean-estimation>

AUTHORS: zengfeng Huang, Ziyue Huang, Yilei WANG, Ke Yi
HIGHLIGHT: We propose a new sparsity-aware algorithm, which improves previous results both theoretically and empirically.

572, TITLE: Global Sparse Momentum SGD for Pruning Very Deep Neural Networks

<https://papers.nips.cc/paper/8867-global-sparse-momentum-sgd-for-pruning-very-deep-neural-networks>

AUTHORS: Xiaohan Ding, guiguang ding, Xiangxin Zhou, Yuchen Guo, Jungong Han, Ji Liu
HIGHLIGHT: In this paper, we propose a novel momentum-SGD-based optimization method to reduce the network complexity by on-the-fly pruning.

573, TITLE: Variational Bayesian Decision-making for Continuous Utilities

<https://papers.nips.cc/paper/8868-variational-bayesian-decision-making-for-continuous-utilities>

AUTHORS: Tomasz Kusmierczyk, Joseph Sakaya, Arto Klami
HIGHLIGHT: We present an automatic pipeline that co-opts continuous utilities into variational inference algorithms to account for decision-making.

574, TITLE: The Normalization Method for Alleviating Pathological Sharpness in Wide Neural Networks

<https://papers.nips.cc/paper/8869-the-normalization-method-for-alleviating-pathological-sharpness-in-wide-neural-networks>

AUTHORS: Ryo Karakida, Shotaro Akaho, Shun-ichi Amari
HIGHLIGHT: We reveal that batch normalization in the last layer contributes to drastically decreasing such pathological sharpness if the width and sample number satisfy a specific condition.

575, TITLE: Single-Model Uncertainties for Deep Learning

<https://papers.nips.cc/paper/8870-single-model-uncertainties-for-deep-learning>

AUTHORS: Natasa Tagasovska, David Lopez-Paz
HIGHLIGHT: To estimate aleatoric uncertainty, we propose Simultaneous Quantile Regression (SQR), a loss function to learn all the conditional quantiles of a given target variable.

576, TITLE: Is Deeper Better only when Shallow is Good?

<https://papers.nips.cc/paper/8871-is-deeper-better-only-when-shallow-is-good>

AUTHORS: Eran Malach, Shai Shalev-Shwartz
HIGHLIGHT: In this work we explore the relation between expressivity properties of deep networks and the ability to train them efficiently using gradient-based algorithms.

577, TITLE: Wasserstein Weisfeiler-Lehman Graph Kernels

<https://papers.nips.cc/paper/8872-wasserstein-weisfeiler-lehman-graph-kernels>

AUTHORS: Matteo Togninalli, Elisabetta Ghisu, Felipe Llinares-Lpez, Bastian Rieck, Karsten Borgwardt
HIGHLIGHT: We propose a novel method that relies on the Wasserstein distance between the node feature vector distributions of two graphs, which allows to find subtler differences in data sets by considering graphs as high-dimensional objects, rather than simple means.

578, TITLE: Domain Generalization via Model-Agnostic Learning of Semantic Features

<https://papers.nips.cc/paper/8873-domain-generalization-via-model-agnostic-learning-of-semantic-features>

AUTHORS: Qi Dou, Daniel Coelho de Castro, Konstantinos Kamnitsas, Ben Glocker
HIGHLIGHT: We investigate the challenging problem of domain generalization, i.e., training a model on multi-domain source data such that it can directly generalize to target domains with unknown statistics.

579, TITLE: Grid Saliency for Context Explanations of Semantic Segmentation

<https://papers.nips.cc/paper/8874-grid-saliency-for-context-explanations-of-semantic-segmentation>

- AUTHORS: Lukas Hoyer, Mauricio Munoz, Prateek Katiyar, Anna Khoreva, Volker Fischer
HIGHLIGHT: To overcome this limitation, we extend the existing approaches to generate grid saliencies, which provide spatially coherent visual explanations for (pixel-level) dense prediction networks.
- 580, TITLE: First-order methods almost always avoid saddle points: The case of vanishing step-sizes
<https://papers.nips.cc/paper/8875-first-order-methods-almost-always-avoid-saddle-points-the-case-of-vanishing-step-sizes>
AUTHORS: Ioannis Panageas, Georgios Piliouras, Xiao Wang
HIGHLIGHT: In this paper, we resolve this question on the affirmative for gradient descent, mirror descent, manifold descent and proximal point.
- 581, TITLE: Maximum Mean Discrepancy Gradient Flow
<https://papers.nips.cc/paper/8876-maximum-mean-discrepancy-gradient-flow>
AUTHORS: Michael Arbel, Anna Korba, Adil SALIM, Arthur Gretton
HIGHLIGHT: We construct a Wasserstein gradient flow of the maximum mean discrepancy (MMD) and study its convergence properties.
- 582, TITLE: Oblivious Sampling Algorithms for Private Data Analysis
<https://papers.nips.cc/paper/8877-oblivious-sampling-algorithms-for-private-data-analysis>
AUTHORS: Sajin Sasy, Olga Ohrimenko
HIGHLIGHT: We study secure and privacy-preserving data analysis based on queries executed on samples from a dataset.
- 583, TITLE: Semi-supervisedly Co-embedding Attributed Networks
<https://papers.nips.cc/paper/8878-semi-supervisedly-co-embedding-attributed-networks>
AUTHORS: Zaiqiao Meng, Shangsong Liang, Jinyuan Fang, Teng Xiao
HIGHLIGHT: In this paper, to deal with the problem, we present a semi-supervised co-embedding model for attributed networks (SCAN) based on the generalized SVAE for the heterogeneous data, which collaboratively learns low-dimensional vector representations of both nodes and attributes for partially labelled attributed networks semi-supervisedly.
- 584, TITLE: From voxels to pixels and back: Self-supervision in natural-image reconstruction from fMRI
<https://papers.nips.cc/paper/8879-from-voxels-to-pixels-and-back-self-supervision-in-natural-image-reconstruction-from-fmri>
AUTHORS: Roman Belyi, Guy Gaziv, Assaf Hoogi, Francesca Strappini, Tal Golan, Michal Irani
HIGHLIGHT: We present a novel approach which, in addition to the scarce labeled data (training pairs), allows to train fMRI-to-image reconstruction networks also on "unlabeled" data (i.e., images without fMRI recording, and fMRI recording without images).
- 585, TITLE: Copulas as High-Dimensional Generative Models: Vine Copula Autoencoders
<https://papers.nips.cc/paper/8880-copulas-as-high-dimensional-generative-models-vine-copula-autoencoders>
AUTHORS: Natasa Tagasovska, Damien Ackerer, Thibault Vatter
HIGHLIGHT: We introduce the vine copula autoencoder (VCAE), a flexible generative model for high-dimensional distributions built in a straightforward three-step procedure.
- 586, TITLE: Nonstochastic Multiarmed Bandits with Unrestricted Delays
<https://papers.nips.cc/paper/8881-nonstochastic-multiarmed-bandits-with-unrestricted-delays>
AUTHORS: Tobias Sommer Thune, Nicolò Cesa-Bianchi, Yevgeny Seldin
HIGHLIGHT: We investigate multiarmed bandits with delayed feedback, where the delays need neither be identical nor bounded.
- 587, TITLE: BIVA: A Very Deep Hierarchy of Latent Variables for Generative Modeling
<https://papers.nips.cc/paper/8882-biva-a-very-deep-hierarchy-of-latent-variables-for-generative-modeling>
AUTHORS: Lars Maaløe, Marco Fraccaro, Valentin Ljubic, Ole Winther
HIGHLIGHT: In this paper we close the performance gap by constructing VAE models that can effectively utilize a deep hierarchy of stochastic variables and model complex covariance structures.
- 588, TITLE: Code Generation as a Dual Task of Code Summarization
<https://papers.nips.cc/paper/8883-code-generation-as-a-dual-task-of-code-summarization>
AUTHORS: Bolin Wei, Ge Li, Xin Xia, Zhiyi Fu, Zhi Jin
HIGHLIGHT: In this paper, we apply the relations between two tasks to improve the performance of both tasks.
- 589, TITLE: Diffeomorphic Temporal Alignment Nets
<https://papers.nips.cc/paper/8884-diffeomorphic-temporal-alignment-nets>
AUTHORS: Ron A. Shapira Weber, Matan Eyal, Nicki Skafté, Oren Shriki, Oren Freifeld

HIGHLIGHT: Here we propose the Diffeomorphic Temporal alignment Net (DTAN), a learning-based method for time-series joint alignment.

590, **TITLE:** Weakly Supervised Instance Segmentation using the Bounding Box Tightness Prior
<https://papers.nips.cc/paper/8885-weakly-supervised-instance-segmentation-using-the-bounding-box-tightness-prior>
AUTHORS: Cheng-Chun Hsu, Kuang-Jui Hsu, Chung-Chi Tsai, Yen-Yu Lin, Yung-Yu Chuang
HIGHLIGHT: This paper presents a weakly supervised instance segmentation method that consumes training data with tight bounding box annotations.

591, **TITLE:** On the Power and Limitations of Random Features for Understanding Neural Networks
<https://papers.nips.cc/paper/8886-on-the-power-and-limitations-of-random-features-for-understanding-neural-networks>
AUTHORS: Gilad Yehudai, Ohad Shamir
HIGHLIGHT: In this paper, we formalize the link between existing results and random features, and argue that despite the impressive positive results, random feature approaches are also inherently limited in what they can explain.

592, **TITLE:** Efficient Pure Exploration in Adaptive Round model
<https://papers.nips.cc/paper/8887-efficient-pure-exploration-in-adaptive-round-model>
AUTHORS: tianyuan jin, Jieming SHI, Xiaokui Xiao, Enhong Chen
HIGHLIGHT: In this paper, we study both PAC and exact top- k arm identification problems and design efficient algorithms considering both round complexity and query complexity.

593, **TITLE:** Multi-objects Generation with Amortized Structural Regularization
<https://papers.nips.cc/paper/8888-multi-objects-generation-with-amortized-structural-regularization>
AUTHORS: Taufik Xu, Chongxuan LI, Jun Zhu, Bo Zhang
HIGHLIGHT: In this paper, we propose amortized structural regularization (ASR), which adopts posterior regularization (PR) to embed human knowledge into DGMs via a set of structural constraints.

594, **TITLE:** Neural Shuffle-Exchange Networks - Sequence Processing in $O(n \log n)$ Time
<https://papers.nips.cc/paper/8889-neural-shuffle-exchange-networks-sequence-processing-in-on-log-n-time>
AUTHORS: Karlis Freivalds, Emils Ozolin?, Agris ?ostaks
HIGHLIGHT: We introduce a new Shuffle-Exchange neural network model for sequence to sequence tasks which have $O(\log n)$ depth and $O(n \log n)$ total complexity.

595, **TITLE:** DetNAS: Backbone Search for Object Detection
<https://papers.nips.cc/paper/8890-detnas-backbone-search-for-object-detection>
AUTHORS: Yukang Chen, Tong Yang, Xiangyu Zhang, GAOFENG MENG, Xinyu Xiao, Jian Sun
HIGHLIGHT: In this work, we present DetNAS to use Neural Architecture Search (NAS) for the design of better backbones for object detection.

596, **TITLE:** Stochastic Proximal Langevin Algorithm: Potential Splitting and Nonasymptotic Rates
<https://papers.nips.cc/paper/8891-stochastic-proximal-langevin-algorithm-potential-splitting-and-nonasymptotic-rates>
AUTHORS: Adil SALIM, Dmitry Koralev, Peter Richtarik
HIGHLIGHT: We propose a new algorithm---Stochastic Proximal Langevin Algorithm (SPLA)---for sampling from a log concave distribution.

597, **TITLE:** Fast AutoAugment
<https://papers.nips.cc/paper/8892-fast-autoaugment>
AUTHORS: Sungbin Lim, Ildoo Kim, Taesup Kim, Chiheon Kim, Sungwoong Kim
HIGHLIGHT: In this paper, we propose an algorithm called Fast AutoAugment that finds effective augmentation policies via a more efficient search strategy based on density matching.

598, **TITLE:** On the Convergence Rate of Training Recurrent Neural Networks
<https://papers.nips.cc/paper/8893-on-the-convergence-rate-of-training-recurrent-neural-networks>
AUTHORS: Zeyuan Allen-Zhu, Yuanzhi Li, Zhao Song
HIGHLIGHT: More importantly, in this paper we build general toolkits to analyze multi-layer networks with ReLU activations.

599, **TITLE:** Interval timing in deep reinforcement learning agents
<https://papers.nips.cc/paper/8894-interval-timing-in-deep-reinforcement-learning-agents>
AUTHORS: Ben Deverett, Ryan Faulkner, Meire Fortunato, Gregory Wayne, Joel Z. Leibo

HIGHLIGHT: In artificial agents, little work has directly addressed (1) which architectural components are necessary for successful development of this ability, (2) how this timing ability comes to be represented in the units and actions of the agent, and (3) whether the resulting behavior of the system converges on solutions similar to those of biology.

600, **TITLE:** Graph-based Discriminators: Sample Complexity and Expressiveness
<https://papers.nips.cc/paper/8895-graph-based-discriminators-sample-complexity-and-expressiveness>
AUTHORS: Roi Livni, Yishay Mansour
HIGHLIGHT: For $k \geq 2$ we introduce a notion similar to the VC-dimension, and show that it controls the sample complexity.

601, **TITLE:** Large Scale Structure of Neural Network Loss Landscapes
<https://papers.nips.cc/paper/8896-large-scale-structure-of-neural-network-loss-landscapes>
AUTHORS: Stanislav Fort, Stanislaw Jastrzebski
HIGHLIGHT: We propose and experimentally verify a unified phenomenological model of the loss landscape that incorporates many of them.

602, **TITLE:** Learning Nonsymmetric Determinantal Point Processes
<https://papers.nips.cc/paper/8897-learning-nonsymmetric-determinantal-point-processes>
AUTHORS: Mike Gartrell, Victor-Emmanuel Brunel, Elvis Dohmatob, Syrine Krichene
HIGHLIGHT: We present a method that enables a tractable algorithm, based on maximum likelihood estimation, for learning nonsymmetric DPPs from data composed of observed subsets.

603, **TITLE:** Hypothesis Set Stability and Generalization
<https://papers.nips.cc/paper/8898-hypothesis-set-stability-and-generalization>
AUTHORS: Dylan J. Foster, Spencer Greenberg, Satyen Kale, Haipeng Luo, Mehryar Mohri, Karthik Sridharan
HIGHLIGHT: We present a study of generalization for data-dependent hypothesis sets.

604, **TITLE:** Learning Object Bounding Boxes for 3D Instance Segmentation on Point Clouds
<https://papers.nips.cc/paper/8899-learning-object-bounding-boxes-for-3d-instance-segmentation-on-point-clouds>
AUTHORS: Bo Yang, Jianan Wang, Ronald Clark, Qingyong Hu, Sen Wang, Andrew Markham, Niki Trigoni
HIGHLIGHT: We propose a novel, conceptually simple and general framework for instance segmentation on 3D point clouds.

605, **TITLE:** Precision-Recall Balanced Topic Modelling
<https://papers.nips.cc/paper/8900-precision-recall-balanced-topic-modelling>
AUTHORS: Seppo Virtanen, Mark Girolami
HIGHLIGHT: We formulate topic modelling as an information retrieval task, where the goal is, based on the latent topic representation, to capture relevant term co-occurrence patterns.

606, **TITLE:** Learning Sparse Distributions using Iterative Hard Thresholding
<https://papers.nips.cc/paper/8901-learning-sparse-distributions-using-iterative-hard-thresholding>
AUTHORS: Jacky Y. Zhang, Rajiv Khanna, Anastasios Kyrillidis, Oluwasanmi O. Koyejo
HIGHLIGHT: In this work, we consider IHT as a solution to the problem of learning sparse discrete distributions.

607, **TITLE:** Discriminative Topic Modeling with Logistic LDA
<https://papers.nips.cc/paper/8902-discriminative-topic-modeling-with-logistic-lda>
AUTHORS: Iryna Korshunova, Hanchen Xiong, Mateusz Fedoryszak, Lucas Theis
HIGHLIGHT: We propose logistic LDA, a novel discriminative variant of latent Dirichlet allocation which is easy to apply to arbitrary inputs.

608, **TITLE:** Quantum Wasserstein Generative Adversarial Networks
<https://papers.nips.cc/paper/8903-quantum-wasserstein-generative-adversarial-networks>
AUTHORS: Shouvanik Chakrabarti, Huang Yiming, Tongyang Li, Soheil Feizi, Xiaodi Wu
HIGHLIGHT: Inspired by previous studies on the adversarial training of classical and quantum generative models, we propose the first design of quantum Wasserstein Generative Adversarial Networks (WGANs), which has been shown to improve the robustness and the scalability of the adversarial training of quantum generative models even on noisy quantum hardware.

609, **TITLE:** Blow: a single-scale hyperconditioned flow for non-parallel raw-audio voice conversion
<https://papers.nips.cc/paper/8904-blow-a-single-scale-hyperconditioned-flow-for-non-parallel-raw-audio-voice-conversion>
AUTHORS: Joan Serra, Santiago Pascual, Carlos Segura Perales

HIGHLIGHT: In this paper, we propose Blow, a single-scale normalizing flow using hypernetwork conditioning to perform many-to-many voice conversion between raw audio.

610, **TITLE:** Hyperparameter Learning via Distributional Transfer
<https://papers.nips.cc/paper/8905-hyperparameter-learning-via-distributional-transfer>
AUTHORS: Ho Chung Law, Peilin Zhao, Leung Sing Chan, Junzhou Huang, Dino Sejdinovic
HIGHLIGHT: We propose to transfer information across tasks using learnt representations of training datasets used in those tasks.

611, **TITLE:** Discriminator optimal transport
<https://papers.nips.cc/paper/8906-discriminator-optimal-transport>
AUTHORS: Akinori Tanaka
HIGHLIGHT: Based on some experiments and a bit of OT theory, we propose discriminator optimal transport (DOT) scheme to improve generated images.

612, **TITLE:** High-dimensional multivariate forecasting with low-rank Gaussian Copula Processes
<https://papers.nips.cc/paper/8907-high-dimensional-multivariate-forecasting-with-low-rank-gaussian-copula-processes>
AUTHORS: David Salinas, Michael Bohlke-Schneider, Laurent Callot, Roberto Medico, Jan Gasthaus
HIGHLIGHT: We propose to combine an RNN-based time series model with a Gaussian copula process output model with a low-rank covariance structure to reduce the computational complexity and handle non-Gaussian marginal distributions.

613, **TITLE:** Are Anchor Points Really Indispensable in Label-Noise Learning?
<https://papers.nips.cc/paper/8908-are-anchor-points-really-indispensable-in-label-noise-learning>
AUTHORS: Xiaobo Xia, Tongliang Liu, Nannan Wang, Bo Han, Chen Gong, Gang Niu, Masashi Sugiyama
HIGHLIGHT: In this paper, without employing anchor points, we propose a transition-revision (T-Revision) method to effectively learn transition matrices, leading to better classifiers.

614, **TITLE:** Aligning Visual Regions and Textual Concepts for Semantic-Grounded Image Representations
<https://papers.nips.cc/paper/8909-aligning-visual-regions-and-textual-concepts-for-semantic-grounded-image-representations>
AUTHORS: Fenglin Liu, Yuanxin Liu, Xuancheng Ren, Xiaodong He, Xu Sun
HIGHLIGHT: In this work, we aim at representing an image with a set of integrated visual regions and corresponding textual concepts, reflecting certain semantics.

615, **TITLE:** Differentiable Ranking and Sorting using Optimal Transport
<https://papers.nips.cc/paper/8910-differentiable-ranking-and-sorting-using-optimal-transport>
AUTHORS: Marco Cuturi, Olivier Teboul, Jean-Philippe Vert
HIGHLIGHT: We propose in this paper to replace the usual sort procedure with a differentiable proxy.

616, **TITLE:** Dichotomize and Generalize: PAC-Bayesian Binary Activated Deep Neural Networks
<https://papers.nips.cc/paper/8911-dichotomize-and-generalize-pac-bayesian-binary-activated-deep-neural-networks>
AUTHORS: Gaël Letarte, Pascal Germain, Benjamin Guedj, Francois Laviolette
HIGHLIGHT: We present a comprehensive study of multilayer neural networks with binary activation, relying on the PAC-Bayesian theory.

617, **TITLE:** Likelihood-Free Overcomplete ICA and Applications In Causal Discovery
<https://papers.nips.cc/paper/8912-likelihood-free-overcomplete-ica-and-applications-in-causal-discovery>
AUTHORS: Chenwei DING, Mingming Gong, Kun Zhang, Dacheng Tao
HIGHLIGHT: To tackle these problems, we present a Likelihood-Free Overcomplete ICA algorithm (LFOICA) that estimates the mixing matrix directly by back-propagation without any explicit assumptions on the density function of independent components.

618, **TITLE:** Interior-Point Methods Strike Back: Solving the Wasserstein Barycenter Problem
<https://papers.nips.cc/paper/8913-interior-point-methods-strike-back-solving-the-wasserstein-barycenter-problem>
AUTHORS: DongDong Ge, Haoyue Wang, Zikai Xiong, Yinyu Ye
HIGHLIGHT: In this paper, we overcome the difficulty by developing a new adapted interior-point method that fully exploits the problem's special matrix structure to reduce the iteration complexity and speed up the Newton procedure.

619, **TITLE:** Beyond Vector Spaces: Compact Data Representation as Differentiable Weighted Graphs
<https://papers.nips.cc/paper/8914-beyond-vector-spaces-compact-data-representation-as-differentiable-weighted-graphs>
AUTHORS: Denis Mazur, Vage Egiazarian, Stanislav Morozov, Artem Babenko
HIGHLIGHT: In this paper, we aim to eliminate the inductive bias imposed by the embedding space geometry.

- 620, TITLE: Subspace Detours: Building Transport Plans that are Optimal on Subspace Projections
<https://papers.nips.cc/paper/8915-subspace-detours-building-transport-plans-that-are-optimal-on-subspace-projections>
AUTHORS: Boris Muzellec, Marco Cuturi
HIGHLIGHT: We propose in this work two methods to extrapolate, from an transport map that is optimal on a subspace, one that is nearly optimal in the entire space.
- 621, TITLE: Efficient Smooth Non-Convex Stochastic Compositional Optimization via Stochastic Recursive Gradient Descent
<https://papers.nips.cc/paper/8916-efficient-smooth-non-convex-stochastic-compositional-optimization-via-stochastic-recursive-gradient-descent>
AUTHORS: Huizhuo Yuan, Xiangru Lian, Chris Junchi Li, Ji Liu, Wenqing Hu
HIGHLIGHT: In this paper, we investigate the stochastic compositional optimization in the general smooth non-convex setting.
- 622, TITLE: On the convergence of single-call stochastic extra-gradient methods
<https://papers.nips.cc/paper/8917-on-the-convergence-of-single-call-stochastic-extra-gradient-methods>
AUTHORS: Yu-Guan Hsieh, Franck Iutzeler, Jérôme Malick, Panayotis Mertikopoulos
HIGHLIGHT: In this paper, we develop a synthetic view of such algorithms, and we complement the existing literature by showing that they retain a $\mathcal{O}(1/t)$ ergodic convergence rate in smooth, deterministic problems.
- 623, TITLE: Infra-slow brain dynamics as a marker for cognitive function and decline
<https://papers.nips.cc/paper/8918-infra-slow-brain-dynamics-as-a-marker-for-cognitive-function-and-decline>
AUTHORS: Shagun Ajmera Shyam Sunder Ajmera, Shreya Rajagopal, Razi Rehman, Devarajan Sridharan
HIGHLIGHT: We investigated this question with a novel application of Gaussian Process Factor Analysis (GPFA) and machine learning to fMRI data.
- 624, TITLE: Robust Principal Component Analysis with Adaptive Neighbors
<https://papers.nips.cc/paper/8919-robust-principal-component-analysis-with-adaptive-neighbors>
AUTHORS: Rui Zhang, Hanghang Tong
HIGHLIGHT: To tackle the issue, we propose a general framework namely robust weight learning with adaptive neighbors (RWL-AN), via which adaptive weight vector is automatically obtained with both robustness and sparse neighbors.
- 625, TITLE: High-Quality Self-Supervised Deep Image Denoising
<https://papers.nips.cc/paper/8920-high-quality-self-supervised-deep-image-denoising>
AUTHORS: Samuli Laine, Tero Karras, Jaakko Lehtinen, Timo Aila
HIGHLIGHT: We describe a novel method for training high-quality image denoising models based on unorganized collections of corrupted images.
- 626, TITLE: Dynamics of stochastic gradient descent for two-layer neural networks in the teacher-student setup
<https://papers.nips.cc/paper/8921-dynamics-of-stochastic-gradient-descent-for-two-layer-neural-networks-in-the-teacher-student-setup>
AUTHORS: Sebastian Goldt, Madhu Advani, Andrew M. Saxe, Florent Krzakala, Lenka Zdeborová
HIGHLIGHT: We show how the dynamics of stochastic gradient descent (SGD) is captured by a set of differential equations and prove that this description is asymptotically exact in the limit of large inputs.
- 627, TITLE: GIFT: Learning Transformation-Invariant Dense Visual Descriptors via Group CNNs
<https://papers.nips.cc/paper/8922-gift-learning-transformation-invariant-dense-visual-descriptors-via-group-cnns>
AUTHORS: Yuan Liu, Zehong Shen, Zhixuan Lin, Sida Peng, Hujun Bao, Xiaowei Zhou
HIGHLIGHT: In this paper, we introduce a novel visual descriptor named Group Invariant Feature Transform (GIFT), which is both discriminative and robust to geometric transformations.
- 628, TITLE: Online Prediction of Switching Graph Labelings with Cluster Specialists
<https://papers.nips.cc/paper/8923-online-prediction-of-switching-graph-labelings-with-cluster-specialists>
AUTHORS: Mark Herbster, James Robinson
HIGHLIGHT: We present an algorithm based on a specialist approach; we develop the machinery of cluster specialists which probabilistically exploits the cluster structure in the graph.
- 629, TITLE: Graph-Based Semi-Supervised Learning with Non-ignorable Non-response
<https://papers.nips.cc/paper/8924-graph-based-semi-supervised-learning-with-non-ignorable-non-response>
AUTHORS: Fan Zhou, Tengfei Li, Haibo Zhou, Hongtu Zhu, Ye Jieping

HIGHLIGHT: To solve the problem, we propose a Graph-based joint model with Non-ignorable Non-response (GNN), followed by a joint inverse weighting estimation procedure incorporated with sampling imputation approach.

630, **TITLE:** BatchBALD: Efficient and Diverse Batch Acquisition for Deep Bayesian Active Learning
<https://papers.nips.cc/paper/8925-batchbald-efficient-and-diverse-batch-acquisition-for-deep-bayesian-active-learning>
AUTHORS: Andreas Kirsch, Joost van Amersfoort, Yariv Gal
HIGHLIGHT: We develop BatchBALD, a tractable approximation to the mutual information between a batch of points and model parameters, which we use as an acquisition function to select multiple informative points jointly for the task of deep Bayesian active learning.

631, **TITLE:** A Mean Field Theory of Quantized Deep Networks: The Quantization-Depth Trade-Off
<https://papers.nips.cc/paper/8926-a-mean-field-theory-of-quantized-deep-networks-the-quantization-depth-trade-off>
AUTHORS: Yaniv Blumenfeld, Dar Gilboa, Daniel Soudry
HIGHLIGHT: We apply mean field techniques to networks with quantized activations in order to evaluate the degree to which quantization degrades signal propagation at initialization.

632, **TITLE:** Beyond Confidence Regions: Tight Bayesian Ambiguity Sets for Robust MDPs
<https://papers.nips.cc/paper/8927-beyond-confidence-regions-tight-bayesian-ambiguity-sets-for-robust-mdps>
AUTHORS: Marek Petrik, Reazul Hasan Russel
HIGHLIGHT: This paper proposes a new paradigm that can achieve better solutions with the same robustness guarantees without using confidence regions as ambiguity sets.

633, **TITLE:** Cross-lingual Language Model Pretraining
<https://papers.nips.cc/paper/8928-cross-lingual-language-model-pretraining>
AUTHORS: Alexis CONNEAU, Guillaume Lample
HIGHLIGHT: In this work, we extend this approach to multiple languages and show the effectiveness of cross-lingual pretraining.

634, **TITLE:** Approximate Bayesian Inference for a Mechanistic Model of Vesicle Release at a Ribbon Synapse
<https://papers.nips.cc/paper/8929-approximate-bayesian-inference-for-a-mechanistic-model-of-vesicle-release-at-a-ribbon-synapse>
AUTHORS: Cornelius Schröder, Ben James, Leon Lagnado, Philipp Berens
HIGHLIGHT: Here, we develop an approximate Bayesian inference scheme for a fully stochastic, biophysically inspired model of glutamate release at the ribbon synapse, a highly specialized synapse found in different sensory systems.

635, **TITLE:** Updates of Equilibrium Prop Match Gradients of Backprop Through Time in an RNN with Static Input
<https://papers.nips.cc/paper/8930-updates-of-equilibrium-prop-match-gradients-of-backprop-through-time-in-an-rnn-with-static-input>
AUTHORS: Maxence Emoult, Benjamin Scellier, Yoshua Bengio, Damien Querlioz, Julie Grollier
HIGHLIGHT: In this work, we introduce a discrete-time version of EP with simplified equations and with reduced simulation time, bringing EP closer to practical machine learning tasks.

636, **TITLE:** Universal Invariant and Equivariant Graph Neural Networks
<https://papers.nips.cc/paper/8931-universal-invariant-and-equivariant-graph-neural-networks>
AUTHORS: Nicolas Keriven, Gabriel Peyr?
HIGHLIGHT: In this paper, we consider a specific class of invariant and equivariant networks, for which we prove new universality theorems.

637, **TITLE:** Are sample means in multi-armed bandits positively or negatively biased?
<https://papers.nips.cc/paper/8932-are-sample-means-in-multi-armed-bandits-positively-or-negatively-biased>
AUTHORS: Jaehyeok Shin, Aaditya Ramdas, Alessandro Rinaldo
HIGHLIGHT: In this paper, we decouple three different sources of this selection bias: adaptive \emph{sampling} of arms, adaptive \emph{stopping} of the experiment, and adaptively \emph{choosing} which arm to study.

638, **TITLE:** On the Correctness and Sample Complexity of Inverse Reinforcement Learning
<https://papers.nips.cc/paper/8933-on-the-correctness-and-sample-complexity-of-inverse-reinforcement-learning>
AUTHORS: Abi Komanduru, Jean Honorio
HIGHLIGHT: A L1-regularized Support Vector Machine formulation of the IRL problem motivated by the geometric analysis is then proposed with the basic objective of the inverse reinforcement problem in mind: to find a reward function that generates a specified optimal policy.

639, **TITLE:** VIREL: A Variational Inference Framework for Reinforcement Learning

<https://papers.nips.cc/paper/8934-virel-a-variational-inference-framework-for-reinforcement-learning>

AUTHORS: Matthew Fellows, Anuj Mahajan, Tim G. J. Rudner, Shimon Whiteson

HIGHLIGHT: We propose VIREL, a theoretically grounded probabilistic inference framework for RL that utilises a parametrised action-value function to summarise future dynamics of the underlying MDP, generalising existing approaches.

640, TITLE: First Order Motion Model for Image Animation

<https://papers.nips.cc/paper/8935-first-order-motion-model-for-image-animation>

AUTHORS: Aliaksandr Siarohin, Stephane Lathuillere, Sergey Tulyakov, Elisa Ricci, Nicu Sebe

HIGHLIGHT: Image animation consists of generating a video sequence so that an object in a source image is animated according to the motion of a driving video. Our framework addresses this problem without using any annotation or prior information about the specific object to animate.

641, TITLE: Tensor Monte Carlo: Particle Methods for the GPU era

<https://papers.nips.cc/paper/8936-tensor-monte-carlo-particle-methods-for-the-gpu-era>

AUTHORS: Laurence Aitchison

HIGHLIGHT: To address these issues, we developed tensor Monte-Carlo (TMC) which gives exponentially many importance samples by separately drawing samples for each of the latent variables, then averaging over all possible combinations.

642, TITLE: Unsupervised Emergence of Egocentric Spatial Structure from Sensorimotor Prediction

<https://papers.nips.cc/paper/8937-unsupervised-emergence-of-egocentric-spatial-structure-from-sensorimotor-prediction>

AUTHORS: Alban Laflaque, Michael Garcia Ortiz

HIGHLIGHT: We propose a simple sensorimotor predictive scheme, apply it to different agents and types of exploration, and evaluate the pertinence of these hypotheses.

643, TITLE: Learning from Label Proportions with Generative Adversarial Networks

<https://papers.nips.cc/paper/8938-learning-from-label-proportions-with-generative-adversarial-networks>

AUTHORS: Jiabin Liu, Bo Wang, Zhiqian Qi, Yingjie Tian, Yong Shi

HIGHLIGHT: In this paper, we leverage generative adversarial networks (GANs) to derive an effective algorithm LLP-GAN for learning from label proportions (LLP), where only the bag-level proportional information in labels is available.

644, TITLE: Efficient and Thrifty Voting by Any Means Necessary

<https://papers.nips.cc/paper/8939-efficient-and-thrifty-voting-by-any-means-necessary>

AUTHORS: Debmalya Mandal, Ariel D. Procaccia, Nisarg Shah, David Woodruff

HIGHLIGHT: We take an unorthodox view of voting by expanding the design space to include both the elicitation rule, whereby voters map their (cardinal) preferences to votes, and the aggregation rule, which transforms the reported votes into collective decisions.

645, TITLE: PointDAN: A Multi-Scale 3D Domain Adaption Network for Point Cloud Representation

<https://papers.nips.cc/paper/8940-pointdan-a-multi-scale-3d-domain-adaption-network-for-point-cloud-representation>

AUTHORS: Can Qin, Haoxuan You, Lichen Wang, C.-C. Jay Kuo, Yun Fu

HIGHLIGHT: In this paper, we propose a novel 3D Domain Adaptation Network for point cloud data (PointDAN).

646, TITLE: ZO-AdaMM: Zeroth-Order Adaptive Momentum Method for Black-Box Optimization

<https://papers.nips.cc/paper/8941-zo-adamm-zeroth-order-adaptive-momentum-method-for-black-box-optimization>

AUTHORS: Xiangyi Chen, Sijia Liu, Kaidi Xu, Xingguo Li, Xue Lin, Mingyi Hong, David Cox

HIGHLIGHT: In this paper, we propose a zeroth-order AdaMM (ZO-AdaMM) algorithm, that generalizes AdaMM to the gradient-free regime.

647, TITLE: Non-Stationary Markov Decision Processes, a Worst-Case Approach using Model-Based Reinforcement Learning

<https://papers.nips.cc/paper/8942-non-stationary-markov-decision-processes-a-worst-case-approach-using-model-based-reinforcement-learning>

AUTHORS: Erwan Lecarpentier, Emmanuel Rachelson

HIGHLIGHT: This work tackles the problem of robust zero-shot planning in non-stationary stochastic environments.

648, TITLE: Depth-First Proof-Number Search with Heuristic Edge Cost and Application to Chemical Synthesis Planning

<https://papers.nips.cc/paper/8943-depth-first-proof-number-search-with-heuristic-edge-cost-and-application-to-chemical-synthesis-planning>

AUTHORS: Akihiro Kishimoto, Beat Buesser, Bei Chen, Adi Botea

HIGHLIGHT: We address this disadvantage of DFPN in RA with a novel approach to combine DFPN with Heuristic Edge Initialization.

- 649, TITLE: Toward a Characterization of Loss Functions for Distribution Learning
<https://papers.nips.cc/paper/8944-toward-a-characterization-of-loss-functions-for-distribution-learning>
AUTHORS: Nika Haghtalab, Cameron Musco, Bo Waggoner
HIGHLIGHT: In this work we study loss functions for learning and evaluating probability distributions over large discrete domains.
- 650, TITLE: Coresets for Archetypal Analysis
<https://papers.nips.cc/paper/8945-coresets-for-archetypal-analysis>
AUTHORS: Sebastian Mair, Ulf Brefeld
HIGHLIGHT: In this paper, we propose efficient coresets for archetypal analysis.
- 651, TITLE: Emergence of Object Segmentation in Perturbed Generative Models
<https://papers.nips.cc/paper/8946-emergence-of-object-segmentation-in-perturbed-generative-models>
AUTHORS: Adam Bielski, Paolo Favaro
HIGHLIGHT: We introduce a novel framework to build a model that can learn how to segment objects from a collection of images without any human annotation.
- 652, TITLE: Optimal Sparse Decision Trees
<https://papers.nips.cc/paper/8947-optimal-sparse-decision-trees>
AUTHORS: Xiyang Hu, Cynthia Rudin, Margo Seltzer
HIGHLIGHT: This work introduces the first practical algorithm for optimal decision trees for binary variables.
- 653, TITLE: Escaping from saddle points on Riemannian manifolds
<https://papers.nips.cc/paper/8948-escaping-from-saddle-points-on-riemannian-manifolds>
AUTHORS: Yue Sun, Nicolas Flammarion, Maryam Fazel
HIGHLIGHT: We consider minimizing a nonconvex, smooth function f on a Riemannian manifold \mathcal{M} .
- 654, TITLE: Multi-source Domain Adaptation for Semantic Segmentation
<https://papers.nips.cc/paper/8949-multi-source-domain-adaptation-for-semantic-segmentation>
AUTHORS: Sicheng Zhao, Bo Li, Xiangyu Yue, Yang Gu, Pengfei Xu, Runbo Hu, Hua Chai, Kurt Keutzer
HIGHLIGHT: In this paper, we propose to investigate multi-source domain adaptation for semantic segmentation.
- 655, TITLE: Localized Structured Prediction
<https://papers.nips.cc/paper/8950-localized-structured-prediction>
AUTHORS: Carlo Ciliberto, Francis Bach, Alessandro Rudi
HIGHLIGHT: In this work we propose the first theoretical framework to deal with part-based data from a general perspective and study a novel method within the setting of statistical learning theory.
- 656, TITLE: Nonzero-sum Adversarial Hypothesis Testing Games
<https://papers.nips.cc/paper/8951-nonzero-sum-adversarial-hypothesis-testing-games>
AUTHORS: Sarath Yasodharan, Patrick Loiseau
HIGHLIGHT: We study nonzero-sum hypothesis testing games that arise in the context of adversarial classification, in both the Bayesian as well as the Neyman-Pearson frameworks.
- 657, TITLE: Manifold-regression to predict from MEG/EEG brain signals without source modeling
<https://papers.nips.cc/paper/8952-manifold-regression-to-predict-from-megeeg-brain-signals-without-source-modeling>
AUTHORS: David Sabbagh, Pierre Ablin, Gael Varoquaux, Alexandre Gramfort, Denis A. Engemann
HIGHLIGHT: In this article, we focus on the task of regression with rank-reduced covariance matrices.
- 658, TITLE: Modeling Tabular data using Conditional GAN
<https://papers.nips.cc/paper/8953-modeling-tabular-data-using-conditional-gan>
AUTHORS: Lei Xu, Maria Skoularidou, Alfredo Cuesta-Infante, Kalyan Veeramachaneni
HIGHLIGHT: We design CTGAN, which uses a conditional generative adversarial network to address these challenges.
- 659, TITLE: Normalization Helps Training of Quantized LSTM
<https://papers.nips.cc/paper/8954-normalization-helps-training-of-quantized-lstm>
AUTHORS: Lu Hou, Jinhua Zhu, James Kwok, Fei Gao, Tao Qin, Tie-Yan Liu

HIGHLIGHT: In this paper, we first show theoretically that training a quantized LSTM is difficult because quantization makes the exploding gradient problem more severe, particularly when the LSTM weight matrices are large. We then show that the popularly used weight/layer/batch normalization schemes can help stabilize the gradient magnitude in training quantized LSTMs.

660, **TITLE:** Trajectory of Alternating Direction Method of Multipliers and Adaptive Acceleration
<https://papers.nips.cc/paper/8955-trajectory-of-alternating-direction-method-of-multipliers-and-adaptive-acceleration>
AUTHORS: Clarice Poon, Jingwei Liang
HIGHLIGHT: By studying the geometric properties of ADMM, we discuss the limitations of current inertial accelerated ADMM and then present and analyze an adaptive acceleration scheme for the method.

661, **TITLE:** Deep Scale-spaces: Equivariance Over Scale
<https://papers.nips.cc/paper/8956-deep-scale-spaces-equivariance-over-scale>
AUTHORS: Daniel Worrall, Max Welling
HIGHLIGHT: We introduce deep scale-spaces, a generalization of convolutional neural networks, exploiting the scale symmetry structure of conventional image recognition tasks.

662, **TITLE:** GRU-ODE-Bayes: Continuous Modeling of Sporadically-Observed Time Series
<https://papers.nips.cc/paper/8957-gru-ode-bayes-continuous-modeling-of-sporadically-observed-time-series>
AUTHORS: Edward De Brouwer, Jaak Simm, Adam Arany, Yves Moreau
HIGHLIGHT: To address these challenges, we propose (1) a continuous-time version of the Gated Recurrent Unit, building upon the recent Neural Ordinary Differential Equations (Chen et al., 2018), and (2) a Bayesian update network that processes the sporadic observations.

663, **TITLE:** Estimating Convergence of Markov chains with L-Lag Couplings
<https://papers.nips.cc/paper/8958-estimating-convergence-of-markov-chains-with-l-lag-couplings>
AUTHORS: Niloy Biswas, Pierre E. Jacob, Paul Vanetti
HIGHLIGHT: We introduce L-lag couplings to generate computable, non-asymptotic upper bound estimates for the total variation or the Wasserstein distance of general Markov chains.

664, **TITLE:** Learning-Based Low-Rank Approximations
<https://papers.nips.cc/paper/8959-learning-based-low-rank-approximations>
AUTHORS: Piotr Indyk, Ali Vakilian, Yang Yuan
HIGHLIGHT: We introduce a "learning-based" algorithm for the low-rank decomposition problem: given an $n \times d$ matrix A , and a parameter k , compute a rank- k matrix A' that minimizes the approximation loss $\|A - A'\|_F$.

665, **TITLE:** Implicit Regularization in Deep Matrix Factorization
<https://papers.nips.cc/paper/8960-implicit-regularization-in-deep-matrix-factorization>
AUTHORS: Sanjeev Arora, Nadav Cohen, Wei Hu, Yuping Luo
HIGHLIGHT: We study the implicit regularization of gradient descent over deep linear neural networks for matrix completion and sensing, a model referred to as deep matrix factorization.

666, **TITLE:** List-decodable Linear Regression
<https://papers.nips.cc/paper/8961-list-decodable-linear-regression>
AUTHORS: Sushrut Karmalkar, Adam Klivans, Pravesh Kothari
HIGHLIGHT: To solve the problem we introduce a new framework for list-decodable learning that strengthens the "identifiability to algorithms" paradigm based on the sum-of-squares method.

667, **TITLE:** Learning elementary structures for 3D shape generation and matching
<https://papers.nips.cc/paper/8962-learning-elementary-structures-for-3d-shape-generation-and-matching>
AUTHORS: Theo Deprelle, Thibault Groueix, Matthew Fisher, Vladimir Kim, Bryan Russell, Mathieu Aubry
HIGHLIGHT: More precisely, we present two complementary approaches to learn elementary structures in a deep learning framework: (i) continuous surface deformation learning and (ii) 3D structure points learning.

668, **TITLE:** On the Hardness of Robust Classification
<https://papers.nips.cc/paper/8963-on-the-hardness-of-robust-classification>
AUTHORS: Pascale Gourdeau, Varun Kanade, Marta Kwiatkowska, James Worrell
HIGHLIGHT: In this paper we study the feasibility of robust learning from the perspective of computational learning theory, considering both sample and computational complexity.

669, **TITLE:** Foundations of Comparison-Based Hierarchical Clustering

<https://papers.nips.cc/paper/8964-foundations-of-comparison-based-hierarchical-clustering>
AUTHORS: Debarghya Ghoshdastidar, Michaël Perrot, Ulrike von Luxburg
HIGHLIGHT: We address the classical problem of hierarchical clustering, but in a framework where one does not have access to a representation of the objects or their pairwise similarities.

670, TITLE: What the Vec? Towards Probabilistically Grounded Embeddings
<https://papers.nips.cc/paper/8965-what-the-vec-towards-probabilistically-grounded-embeddings>
AUTHORS: Carl Allen, Ivana Balazevic, Timothy Hospedales
HIGHLIGHT: We show that different interactions of PMI vectors encode semantic properties that can be captured in low dimensional word embeddings by suitable projection, theoretically explaining why the embeddings of W2V and Glove work, and, in turn, revealing an interesting mathematical interconnection between the semantic relationships of relatedness, similarity, paraphrase and analogy.

671, TITLE: Minimizers of the Empirical Risk and Risk Monotonicity
<https://papers.nips.cc/paper/8966-minimizers-of-the-empirical-risk-and-risk-monotonicity>
AUTHORS: Marco Loog, Tom Viering, Alexander Mey
HIGHLIGHT: Our work introduces the formal notion of risk monotonicity, which asks the risk to not deteriorate with increasing training set sizes in expectation over the training samples.

672, TITLE: Explicit Planning for Efficient Exploration in Reinforcement Learning
<https://papers.nips.cc/paper/8967-explicit-planning-for-efficient-exploration-in-reinforcement-learning>
AUTHORS: Liangpeng Zhang, Ke Tang, Xin Yao
HIGHLIGHT: We argue that explicit planning for exploration can help alleviate such a problem, and propose a Value Iteration for Exploration Cost (VIEC) algorithm which computes the optimal exploration scheme by solving an augmented MDP.

673, TITLE: Lower Bounds on Adversarial Robustness from Optimal Transport
<https://papers.nips.cc/paper/8968-lower-bounds-on-adversarial-robustness-from-optimal-transport>
AUTHORS: Arjun Nitin Bhagoji, Daniel Cullina, Prateek Mittal
HIGHLIGHT: In this paper, we use optimal transport to characterize the maximum achievable accuracy in an adversarial classification scenario.

674, TITLE: Neural Spline Flows
<https://papers.nips.cc/paper/8969-neural-spline-flows>
AUTHORS: Conor Durkan, Artur Bekasov, Iain Murray, George Papamakarios
HIGHLIGHT: Building upon recent work, we propose a fully-differentiable module based on monotonic rational-quadratic splines, which enhances the flexibility of both coupling and autoregressive transforms while retaining analytic invertibility.

675, TITLE: Phase Transitions and Cyclic Phenomena in Bandits with Switching Constraints
<https://papers.nips.cc/paper/8970-phase-transitions-and-cyclic-phenomena-in-bandits-with-switching-constraints>
AUTHORS: David Simchi-Levi, Yunzong Xu
HIGHLIGHT: We consider the classical stochastic multi-armed bandit problem with a constraint on the total cost incurred by switching between actions.

676, TITLE: Latent Weights Do Not Exist: Rethinking Binarized Neural Network Optimization
<https://papers.nips.cc/paper/8971-latent-weights-do-not-exist-rethinking-binarized-neural-network-optimization>
AUTHORS: Koen Helwegen, James Widdicombe, Lukas Geiger, Zechun Liu, Kwang-Ting Cheng, Roeland Nusselder
HIGHLIGHT: In this paper, we argue that these latent weights cannot be treated analogously to weights in real-valued networks.

677, TITLE: Nonlinear scaling of resource allocation in sensory bottlenecks
<https://papers.nips.cc/paper/8972-nonlinear-scaling-of-resource-allocation-in-sensory-bottlenecks>
AUTHORS: Laura Rose Edmondson, Alejandro Jimenez Rodriguez, Hannes P. Saal
HIGHLIGHT: Here, we show analytically and numerically that resource allocation scales nonlinearly in efficient coding models that maximize information transfer, when inputs arise from separate regions with different receptor densities.

678, TITLE: Constrained Reinforcement Learning Has Zero Duality Gap
<https://papers.nips.cc/paper/8973-constrained-reinforcement-learning-has-zero-duality-gap>
AUTHORS: Santiago Paternain, Luiz Chamon, Miguel Calvo-Fullana, Alejandro Ribeiro
HIGHLIGHT: This work provides theoretical support to these approaches by establishing that despite its non-convexity, this problem has zero duality gap, i.e., it can be solved exactly in the dual domain, where it becomes convex.

- 679, TITLE: Symmetry-adapted generation of 3d point sets for the targeted discovery of molecules
<https://papers.nips.cc/paper/8974-symmetry-adapted-generation-of-3d-point-sets-for-the-targeted-discovery-of-molecules>
AUTHORS: Niklas Gebauer, Michael Gastegger, Kristof Sch?tt
HIGHLIGHT: Here, we introduce a generative neural network for 3d point sets that respects the rotational invariance of the targeted structures.
- 680, TITLE: An adaptive nearest neighbor rule for classification
<https://papers.nips.cc/paper/8975-an-adaptive-nearest-neighbor-rule-for-classification>
AUTHORS: Akshay Balsubramani, Sanjoy Dasgupta, yoav Freund, Shay Moran
HIGHLIGHT: We introduce a variant of the k -nearest neighbor classifier in which k is chosen adaptively for each query, rather than supplied as a parameter.
- 681, TITLE: Coresets for Clustering with Fairness Constraints
<https://papers.nips.cc/paper/8976-coresets-for-clustering-with-fairness-constraints>
AUTHORS: Lingxiao Huang, Shaofeng Jiang, Nisheeth Vishnoi
HIGHLIGHT: The main contribution of this paper is an approach to clustering with fairness constraints that involve ℓ multiple, non-disjoint attributes, that is ℓ also scalable}.
- 682, TITLE: PerspectiveNet: A Scene-consistent Image Generator for New View Synthesis in Real Indoor Environments
<https://papers.nips.cc/paper/8977-perspectivenet-a-scene-consistent-image-generator-for-new-view-synthesis-in-real-indoor-environments>
AUTHORS: Ben Graham, David Novotny, Jeremy Reizenstein
HIGHLIGHT: Given a set of a reference RGBD views of an indoor environment, and a new viewpoint, our goal is to predict the view from that location.
- 683, TITLE: MAVEN: Multi-Agent Variational Exploration
<https://papers.nips.cc/paper/8978-maven-multi-agent-variational-exploration>
AUTHORS: Anuj Mahajan, Tabish Rashid, Mikayel Samvelyan, Shimon Whiteson
HIGHLIGHT: In this paper, we analyse value-based methods that are known to have superior performance in complex environments.
- 684, TITLE: Competitive Gradient Descent
<https://papers.nips.cc/paper/8979-competitive-gradient-descent>
AUTHORS: Florian Schaefer, Anima Anandkumar
HIGHLIGHT: We introduce a new algorithm for the numerical computation of Nash equilibria of competitive two-player games.
- 685, TITLE: Globally Convergent Newton Methods for Ill-conditioned Generalized Self-concordant Losses
<https://papers.nips.cc/paper/8980-globally-convergent-newton-methods-for-ill-conditioned-generalized-self-concordant-losses>
AUTHORS: Ulysse Marteau-Ferey, Francis Bach, Alessandro Rudi
HIGHLIGHT: In this paper, we study large-scale convex optimization algorithms based on the Newton method applied to regularized generalized self-concordant losses, which include logistic regression and softmax regression.
- 686, TITLE: Continual Unsupervised Representation Learning
<https://papers.nips.cc/paper/8981-continual-unsupervised-representation-learning>
AUTHORS: Dushyant Rao, Francesco Visin, Andrei Rusu, Razvan Pascanu, Yee Whye Teh, Raia Hadsell
HIGHLIGHT: In this work, we propose an approach (CURL) to tackle a more general problem that we will refer to as unsupervised continual learning.
- 687, TITLE: Self-Routing Capsule Networks
<https://papers.nips.cc/paper/8982-self-routing-capsule-networks>
AUTHORS: Taeyoung Hahn, Myeongjang Pyeon, Gunhee Kim
HIGHLIGHT: In this work, we propose a novel and surprisingly simple routing strategy called self-routing where each capsule is routed independently by its subordinate routing network.
- 688, TITLE: The Parameterized Complexity of Cascading Portfolio Scheduling
<https://papers.nips.cc/paper/8983-the-parameterized-complexity-of-cascading-portfolio-scheduling>
AUTHORS: Eduard Eiben, Robert Galian, Iyad Kanj, Stefan Szeider
HIGHLIGHT: In this paper we study the parameterized complexity of this problem and establish its fixed-parameter tractability by utilizing structural properties of the success relation between algorithms and test instances.

- 689, TITLE: Maximum Expected Hitting Cost of a Markov Decision Process and Informativeness of Rewards
<https://papers.nips.cc/paper/8984-maximum-expected-hitting-cost-of-a-markov-decision-process-and-informativeness-of-rewards>
AUTHORS: Falcon Dai, Matthew Walter
HIGHLIGHT: We propose a new complexity measure for Markov decision processes (MDPs), the maximum expected hitting cost (MEHC).
- 690, TITLE: Bipartite expander Hopfield networks as self-decoding high-capacity error correcting codes
<https://papers.nips.cc/paper/8985-bipartite-expander-hopfield-networks-as-self-decoding-high-capacity-error-correcting-codes>
AUTHORS: Rishidev Chaudhuri, Ila Fiete
HIGHLIGHT: We prove that it is possible to construct an associative content-addressable network that combines the properties of strong error correcting codes and Hopfield networks: it simultaneously possesses exponentially many stable states, these states are robust enough, with large enough basins of attraction that they can be correctly recovered despite errors in a finite fraction of all nodes, and the errors are intrinsically corrected by the network's own dynamics.
- 691, TITLE: Sequence Modeling with Unconstrained Generation Order
<https://papers.nips.cc/paper/8986-sequence-modeling-with-unconstrained-generation-order>
AUTHORS: Dmitrii Emelianenko, Elena Voita, Pavel Serdyukov
HIGHLIGHT: In contrast, we propose a more general model that can generate the output sequence by inserting tokens in any arbitrary order.
- 692, TITLE: Probabilistic Logic Neural Networks for Reasoning
<https://papers.nips.cc/paper/8987-probabilistic-logic-neural-networks-for-reasoning>
AUTHORS: Meng Qu, Jian Tang
HIGHLIGHT: In this paper, we propose the probabilistic Logic Neural Network (pLogicNet), which combines the advantages of both methods.
- 693, TITLE: A Polynomial Time Algorithm for Log-Concave Maximum Likelihood via Locally Exponential Families
<https://papers.nips.cc/paper/8988-a-polynomial-time-algorithm-for-log-concave-maximum-likelihood-via-locally-exponential-families>
AUTHORS: Brian Axelrod, Ilias Diakonikolas, Alistair Stewart, Anastasios Sidiropoulos, Gregory Valiant
HIGHLIGHT: Specifically, we present an algorithm which, given n points in \mathbb{R}^d and an accuracy parameter $\epsilon > 0$, runs in time $\text{poly}(n, d, 1/\epsilon)$, and returns a log-concave distribution which, with high probability, has the property that the likelihood of the n points under the returned distribution is at most an additive ϵ less than the maximum likelihood that could be achieved via any log-concave distribution.
- 694, TITLE: A Unifying Framework for Spectrum-Preserving Graph Sparsification and Coarsening
<https://papers.nips.cc/paper/8989-a-unifying-framework-for-spectrum-preserving-graph-sparsification-and-coarsening>
AUTHORS: GeCIA Bravo Hermsdorff, Lee Gunderson
HIGHLIGHT: In this work, we provide a unifying framework that captures both of these operations, allowing one to simultaneously sparsify and coarsen a graph while preserving its large-scale structure.
- 695, TITLE: Stochastic Runge-Kutta Accelerates Langevin Monte Carlo and Beyond
<https://papers.nips.cc/paper/8990-stochastic-runge-kutta-accelerates-langevin-monte-carlo-and-beyond>
AUTHORS: Xuechen Li, Yi Wu, Lester Mackey
HIGHLIGHT: In this paper, we establish the convergence rate of sampling algorithms obtained by discretizing smooth Itô diffusions exhibiting fast Wasserstein contraction, based on local deviation properties of the integration scheme.
- 696, TITLE: The Implicit Bias of AdaGrad on Separable Data
<https://papers.nips.cc/paper/8991-the-implicit-bias-of-adagrad-on-separable-data>
AUTHORS: Qian Qian, Xiaoyuan Qian
HIGHLIGHT: We study the implicit bias of AdaGrad on separable linear classification problems.
- 697, TITLE: On two ways to use determinantal point processes for Monte Carlo integration
<https://papers.nips.cc/paper/8992-on-two-ways-to-use-determinantal-point-processes-for-monte-carlo-integration>
AUTHORS: Guillaume Gautier, Rémi Bardenet, Michal Valko
HIGHLIGHT: In this paper, we first take the EZ estimator out of the cellar, and analyze it using modern arguments. Second, we provide an efficient implementation to sample exactly a particular multidimensional DPP called multivariate Jacobi ensemble.
- 698, TITLE: LiteEval: A Coarse-to-Fine Framework for Resource Efficient Video Recognition
<https://papers.nips.cc/paper/8993-liteeval-a-coarse-to-fine-framework-for-resource-efficient-video-recognition>

AUTHORS: Zuxuan Wu, Caiming Xiong, Yu-Gang Jiang, Larry S. Davis
HIGHLIGHT: This paper presents LiteEval, a simple yet effective coarse-to-fine framework for resource efficient video recognition, suitable for both online and offline scenarios.

699, TITLE: How degenerate is the parametrization of neural networks with the ReLU activation function?
<https://papers.nips.cc/paper/8994-how-degenerate-is-the-parametrization-of-neural-networks-with-the-relu-activation-function>
AUTHORS: Dennis Maximilian Elbr?chter, Julius Berner, Philipp Grohs
HIGHLIGHT: We present pathologies which prevent inverse stability in general, and, for shallow networks, proceed to establish a restricted space of parametrizations on which we have inverse stability w.r.t. to a Sobolev norm.

700, TITLE: Spike-Train Level Backpropagation for Training Deep Recurrent Spiking Neural Networks
<https://papers.nips.cc/paper/8995-spike-train-level-backpropagation-for-training-deep-recurrent-spiking-neural-networks>
AUTHORS: Wenrui Zhang, Peng Li
HIGHLIGHT: To enable supervised training of RSNNs under a well-defined loss function, we present a novel Spike-Train level RSNNs Backpropagation (ST-RSBP) algorithm for training deep RSNNs.

701, TITLE: Re-examination of the Role of Latent Variables in Sequence Modeling
<https://papers.nips.cc/paper/8996-re-examination-of-the-role-of-latent-variables-in-sequence-modeling>
AUTHORS: Guokun Lai, Zihang Dai, Yiming Yang, Shinjae Yoo
HIGHLIGHT: Our analysis reveals that under the restriction of fully factorized output distribution in previous evaluations, the stochastic variants were implicitly leveraging intra-step correlation but the deterministic recurrent baselines were prohibited to do so, resulting in an unfair comparison.

702, TITLE: Max-value Entropy Search for Multi-Objective Bayesian Optimization
<https://papers.nips.cc/paper/8997-max-value-entropy-search-for-multi-objective-bayesian-optimization>
AUTHORS: Syrine Belakaria, Aryan Deshwal, Janardhan Rao Doppa
HIGHLIGHT: We propose a novel approach referred to as Max-value Entropy Search for Multi-objective Optimization (MESMO) to solve this problem.

703, TITLE: Stein Variational Gradient Descent With Matrix-Valued Kernels
<https://papers.nips.cc/paper/8998-stein-variational-gradient-descent-with-matrix-valued-kernels>
AUTHORS: Dilin Wang, Ziyang Tang, Chandrajit Bajaj, Qiang Liu
HIGHLIGHT: In this work, we enhance SVGD by leveraging preconditioning matrices, such as the Hessian and Fisher information matrix, to incorporate geometric information into SVGD updates.

704, TITLE: Crowdsourcing via Pairwise Co-occurrences: Identifiability and Algorithms
<https://papers.nips.cc/paper/8999-crowdsourcing-via-pairwise-co-occurrences-identifiability-and-algorithms>
AUTHORS: Shahana Ibrahim, Xiao Fu, Nikolaos Kargas, Kejun Huang
HIGHLIGHT: We propose an algebraic algorithm reminiscent of convex geometry-based structured matrix factorization to solve the model identification problem efficiently, and an identifiability-enhanced algorithm for handling more challenging and critical scenarios.

705, TITLE: Detecting Overfitting via Adversarial Examples
<https://papers.nips.cc/paper/9000-detecting-overfitting-via-adversarial-examples>
AUTHORS: Roman Werpachowski, Andr?s Gy?rgy, Csaba Szepesvari
HIGHLIGHT: We propose a new hypothesis test that uses only the original test data to detect overfitting.

706, TITLE: A Unified Bellman Optimality Principle Combining Reward Maximization and Empowerment
<https://papers.nips.cc/paper/9001-a-unified-bellman-optimality-principle-combining-reward-maximization-and-empowerment>
AUTHORS: Felix Leibfried, Sergio Pascual-D?az, Jordi Grau-Moya
HIGHLIGHT: In this paper, we investigate the use of empowerment in the presence of an extrinsic reward signal.

707, TITLE: SMILE: Scalable Meta Inverse Reinforcement Learning through Context-Conditional Policies
<https://papers.nips.cc/paper/9002-smile-scalable-meta-inverse-reinforcement-learning-through-context-conditional-policies>
AUTHORS: Seyed Kamyar Seyed Ghasemipour, Shixiang (Shane) Gu, Richard Zemel
HIGHLIGHT: In this work, we propose SMILE, a scalable framework for Meta Inverse Reinforcement Learning (Meta-IRL) based on maximum entropy IRL, which can learn high-quality policies from few demonstrations.

708, TITLE: Towards Understanding the Importance of Shortcut Connections in Residual Networks
<https://papers.nips.cc/paper/9003-towards-understanding-the-importance-of-shortcut-connections-in-residual-networks>

AUTHORS: Tianyi Liu, Minshuo Chen, Mo Zhou, Simon S. Du, Enlu Zhou, Tuo Zhao
HIGHLIGHT: In this paper, we study a two-layer non-overlapping convolutional ResNet.

709, TITLE: Modular Universal Reparameterization: Deep Multi-task Learning Across Diverse Domains
<https://papers.nips.cc/paper/9004-modular-universal-reparameterization-deep-multi-task-learning-across-diverse-domains>
AUTHORS: Elliot Meyerson, Risto Miikkulainen
HIGHLIGHT: To approach this question, deep multi-task learning is extended in this paper to the setting where there is no obvious overlap between task architectures.

710, TITLE: Solving Interpretable Kernel Dimensionality Reduction
<https://papers.nips.cc/paper/9005-solving-interpretable-kernel-dimensionality-reduction>
AUTHORS: Chieh Wu, Jared Miller, Yale Chang, Mario Sznajder, Jennifer Dy
HIGHLIGHT: This work extends the theoretical guarantees of ISM to an entire family of kernels, thereby empowering ISM to solve any kernel method of the same objective.

711, TITLE: Interaction Hard Thresholding: Consistent Sparse Quadratic Regression in Sub-quadratic Time and Space
<https://papers.nips.cc/paper/9006-interaction-hard-thresholding-consistent-sparse-quadratic-regression-in-sub-quadratic-time-and-space>
AUTHORS: Shuo Yang, Yanyao Shen, Sujay Sanghavi
HIGHLIGHT: In this paper, we provide a new algorithm - Interaction Hard Thresholding (IntHT) which is the first one to provably accurately solve this problem in sub-quadratic time and space.

712, TITLE: A Model to Search for Synthesizable Molecules
<https://papers.nips.cc/paper/9007-a-model-to-search-for-synthesizable-molecules>
AUTHORS: John Bradshaw, Brooks Paige, Matt J. Kusner, Marwin Segler, Jos? Miguel Hernandez-Lobato
HIGHLIGHT: We propose a new molecule generation model, mirroring a more realistic real-world process, where (a) reactants are selected, and (b) combined to form more complex molecules.

713, TITLE: Post training 4-bit quantization of convolutional networks for rapid-deployment
<https://papers.nips.cc/paper/9008-post-training-4-bit-quantization-of-convolutional-networks-for-rapid-deployment>
AUTHORS: Ron Banner, Yury Nahshan, Daniel Soudry
HIGHLIGHT: This paper introduces the first practical 4-bit post training quantization approach: it does not involve training the quantized model (fine-tuning), nor it requires the availability of the full dataset.

714, TITLE: Fast and Flexible Multi-Task Classification using Conditional Neural Adaptive Processes
<https://papers.nips.cc/paper/9009-fast-and-flexible-multi-task-classification-using-conditional-neural-adaptive-processes>
AUTHORS: James Requeima, Jonathan Gordon, John Bronskill, Sebastian Nowozin, Richard E. Turner
HIGHLIGHT: The goal of this paper is to design image classification systems that, after an initial multi-task training phase, can automatically adapt to new tasks encountered at test time.

715, TITLE: Differentially Private Anonymized Histograms
<https://papers.nips.cc/paper/9010-differentially-private-anonymized-histograms>
AUTHORS: Ananda Theertha Suresh
HIGHLIGHT: Motivated by these applications, we propose the first differentially private mechanism to release anonymized histograms that achieves near-optimal privacy utility trade-off both in terms of number of items and the privacy parameter.

716, TITLE: Dynamic Local Regret for Non-convex Online Forecasting
<https://papers.nips.cc/paper/9011-dynamic-local-regret-for-non-convex-online-forecasting>
AUTHORS: Sergul Aydore, Tianhao Zhu, Dean P. Foster
HIGHLIGHT: We introduce a local regret for non-convex models in a dynamic environment.

717, TITLE: Learning Local Search Heuristics for Boolean Satisfiability
<https://papers.nips.cc/paper/9012-learning-local-search-heuristics-for-boolean-satisfiability>
AUTHORS: Emre Yolcu, Barnabas Póczos
HIGHLIGHT: We present an approach to learn SAT solver heuristics from scratch through deep reinforcement learning with a curriculum.

718, TITLE: Provably Efficient Q-Learning with Low Switching Cost
<https://papers.nips.cc/paper/9013-provably-efficient-q-learning-with-low-switching-cost>
AUTHORS: Yu Bai, Tengyang Xie, Nan Jiang, Yu-Xiang Wang

HIGHLIGHT: Our main contribution, Q-Learning with UCB2 exploration, is a model-free algorithm for H -step episodic MDP that achieves sublinear regret whose local switching cost in K episodes is $O(H^3SA \log K)$, and we provide a lower bound of $\Omega(HSA)$ on the local switching cost for any no-regret algorithm.

719, **TITLE:** Solving graph compression via optimal transport
<https://papers.nips.cc/paper/9014-solving-graph-compression-via-optimal-transport>
AUTHORS: Vikas Garg, Tommi Jaakkola
HIGHLIGHT: We propose a new approach to graph compression by appeal to optimal transport.

720, **TITLE:** PyTorch: An Imperative Style, High-Performance Deep Learning Library
<https://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library>
AUTHORS: Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, Soumith Chintala
HIGHLIGHT: In this paper, we detail the principles that drove the implementation of PyTorch and how they are reflected in its architecture.

721, **TITLE:** Stability of Graph Scattering Transforms
<https://papers.nips.cc/paper/9016-stability-of-graph-scattering-transforms>
AUTHORS: Fernando Gama, Alejandro Ribeiro, Joan Bruna
HIGHLIGHT: In this work, we extend scattering transforms to network data by using multi-resolution graph wavelets, whose computation can be obtained by means of graph convolutions.

722, **TITLE:** A Debiased MDI Feature Importance Measure for Random Forests
<https://papers.nips.cc/paper/9017-a-debiased-mdi-feature-importance-measure-for-random-forests>
AUTHORS: Xiao Li, Yu Wang, Sumanta Basu, Karl Kumbier, Bin Yu
HIGHLIGHT: In this paper, we address the feature selection bias of MDI from both theoretical and methodological perspectives.

723, **TITLE:** Provably Efficient Q-learning with Function Approximation via Distribution Shift Error Checking Oracle
<https://papers.nips.cc/paper/9018-provably-efficient-q-learning-with-function-approximation-via-distribution-shift-error-checking-oracle>
AUTHORS: Simon S. Du, Yuping Luo, Ruosong Wang, Hanrui Zhang
HIGHLIGHT: The current paper presents a provably efficient algorithm for Q-learning with linear function approximation.

724, **TITLE:** Sparse Logistic Regression Learns All Discrete Pairwise Graphical Models
<https://papers.nips.cc/paper/9019-sparse-logistic-regression-learns-all-discrete-pairwise-graphical-models>
AUTHORS: Shanshan Wu, Sujay Sanghavi, Alexandros G. Dimakis
HIGHLIGHT: The algorithm is (appropriately regularized) maximum conditional log-likelihood, which involves solving a convex program for each node; for Ising models this is ℓ_1 -constrained logistic regression, while for more general alphabets an group-norm constraint needs to be used.

725, **TITLE:** Fast Convergence of Natural Gradient Descent for Over-Parameterized Neural Networks
<https://papers.nips.cc/paper/9020-fast-convergence-of-natural-gradient-descent-for-over-parameterized-neural-networks>
AUTHORS: Guodong Zhang, James Martens, Roger B. Grosse
HIGHLIGHT: In this work, we analyze for the first time the speed of convergence to global optimum for natural gradient descent on non-linear neural networks with the squared error loss.

726, **TITLE:** Rapid Convergence of the Unadjusted Langevin Algorithm: Isoperimetry Suffices
<https://papers.nips.cc/paper/9021-rapid-convergence-of-the-unadjusted-langevin-algorithm-isoperimetry-suffices>
AUTHORS: Santosh Vempala, Andre Wibisono
HIGHLIGHT: We study the Unadjusted Langevin Algorithm (ULA) for sampling from a probability distribution $\nu = e^{-f}$ on \mathbb{R}^n .

727, **TITLE:** Learning Distributions Generated by One-Layer ReLU Networks
<https://papers.nips.cc/paper/9022-learning-distributions-generated-by-one-layer-relu-networks>
AUTHORS: Shanshan Wu, Alexandros G. Dimakis, Sujay Sanghavi
HIGHLIGHT: We consider the problem of estimating the parameters of a d -dimensional rectified Gaussian distribution from i.i.d. samples.

- 728, TITLE: Large-scale optimal transport map estimation using projection pursuit
<https://papers.nips.cc/paper/9023-large-scale-optimal-transport-map-estimation-using-projection-pursuit>
AUTHORS: Cheng Meng, Yuan Ke, Jingyi Zhang, Mengrui Zhang, Wenxuan Zhong, Ping Ma
HIGHLIGHT: Instead, we propose an estimation method of large-scale OTM by combining the idea of projection pursuit regression and sufficient dimension reduction.
- 729, TITLE: A Structured Prediction Approach for Generalization in Cooperative Multi-Agent Reinforcement Learning
<https://papers.nips.cc/paper/9024-a-structured-prediction-approach-for-generalization-in-cooperative-multi-agent-reinforcement-learning>
AUTHORS: Nicolas Carion, Nicolas Usunier, Gabriel Synnaeve, Alessandro Lazaric
HIGHLIGHT: By leveraging this property, we introduce a novel structured prediction approach to assign agents to tasks.
- 730, TITLE: On Exact Computation with an Infinitely Wide Neural Net
<https://papers.nips.cc/paper/9025-on-exact-computation-with-an-infinitely-wide-neural-net>
AUTHORS: Sanjeev Arora, Simon S. Du, Wei Hu, Zhiyuan Li, Russ R. Salakhutdinov, Ruosong Wang
HIGHLIGHT: The current paper gives the first efficient exact algorithm for computing the extension of NTK to convolutional neural nets, which we call Convolutional NTK (CNTK), as well as an efficient GPU implementation of this algorithm.
- 731, TITLE: Loaded DiCE: Trading off Bias and Variance in Any-Order Score Function Gradient Estimators for Reinforcement Learning
<https://papers.nips.cc/paper/9026-loaded-dice-trading-off-bias-and-variance-in-any-order-score-function-gradient-estimators-for-reinforcement-learning>
AUTHORS: Gregory Farquhar, Shimon Whiteson, Jakob Foerster
HIGHLIGHT: Our objective is compatible with arbitrary advantage estimators, which allows the control of the bias and variance of any-order derivatives when using function approximation.
- 732, TITLE: Chirality Nets for Human Pose Regression
<https://papers.nips.cc/paper/9027-chirality-nets-for-human-pose-regression>
AUTHORS: Raymond Yeh, Yuan-Ting Hu, Alexander Schwing
HIGHLIGHT: We propose Chirality Nets, a family of deep nets that is equivariant to the "chirality transform," i.e., the transformation to create a chiral pair.
- 733, TITLE: Efficient Approximation of Deep ReLU Networks for Functions on Low Dimensional Manifolds
<https://papers.nips.cc/paper/9028-efficient-approximation-of-deep-relu-networks-for-functions-on-low-dimensional-manifolds>
AUTHORS: Minshuo Chen, Haoming Jiang, Wenjing Liao, Tuo Zhao
HIGHLIGHT: In this paper, we prove that neural networks can efficiently approximate functions supported on low dimensional manifolds.
- 734, TITLE: Fast Decomposable Submodular Function Minimization using Constrained Total Variation
<https://papers.nips.cc/paper/9029-fast-decomposable-submodular-function-minimization-using-constrained-total-variation>
AUTHORS: Senanayak Sesh Kumar Karri, Francis Bach, Thomas Pock
HIGHLIGHT: In this paper, we consider a modified convex problem requiring constrained version of the total variation oracles that can be solved with significantly fewer calls to the simple minimization oracles.
- 735, TITLE: Which Algorithmic Choices Matter at Which Batch Sizes? Insights From a Noisy Quadratic Model
<https://papers.nips.cc/paper/9030-which-algorithmic-choices-matter-at-which-batch-sizes-insights-from-a-noisy-quadratic-model>
AUTHORS: Guodong Zhang, Lala Li, Zachary Nado, James Martens, Sushant Sachdeva, George Dahl, Chris Shallue, Roger B. Grosse
HIGHLIGHT: In this work, we study how the critical batch size changes based on properties of the optimization algorithm, including acceleration and preconditioning, through two different lenses: large scale experiments and analysis using a simple noisy quadratic model (NQM).
- 736, TITLE: Spherical Text Embedding
<https://papers.nips.cc/paper/9031-spherical-text-embedding>
AUTHORS: Yu Meng, Jiabin Huang, Guangyuan Wang, Chao Zhang, Honglei Zhuang, Lance Kaplan, Jiawei Han
HIGHLIGHT: To close this gap, we propose a spherical generative model based on which unsupervised word and paragraph embeddings are jointly learned.
- 737, TITLE: Mobius Transformation for Fast Inner Product Search on Graph
<https://papers.nips.cc/paper/9032-mobius-transformation-for-fast-inner-product-search-on-graph>
AUTHORS: Zhixin Zhou, Shulong Tan, Zhaozhuo Xu, Ping Li

- HIGHLIGHT: We present a fast search on graph algorithm for Maximum Inner Product Search (MIPS).
- 738, TITLE: Hyperbolic Graph Neural Networks
<https://papers.nips.cc/paper/9033-hyperbolic-graph-neural-networks>
AUTHORS: Qi Liu, Maximilian Nickel, Douwe Kiela
HIGHLIGHT: Motivated by recent advances in geometric representation learning, we propose a novel GNN architecture for learning representations on Riemannian manifolds with differentiable exponential and logarithmic maps.
- 739, TITLE: Average Individual Fairness: Algorithms, Generalization and Experiments
<https://papers.nips.cc/paper/9034-average-individual-fairness-algorithms-generalization-and-experiments>
AUTHORS: Saeed Sharifi-Malvajerdi, Michael Kearns, Aaron Roth
HIGHLIGHT: We propose a new family of fairness definitions for classification problems that combine some of the best properties of both statistical and individual notions of fairness.
- 740, TITLE: Fixing the train-test resolution discrepancy
<https://papers.nips.cc/paper/9035-fixing-the-train-test-resolution-discrepancy>
AUTHORS: Hugo Touvron, Andrea Vedaldi, Matthijs Douze, Herve Jegou
HIGHLIGHT: This paper first shows that existing augmentations induce a significant discrepancy between the size of the objects seen by the classifier at train and test time: in fact, a lower train resolution improves the classification at test time! We then propose a simple strategy to optimize the classifier performance, that employs different train and test resolutions.
- 741, TITLE: Modeling Dynamic Functional Connectivity with Latent Factor Gaussian Processes
<https://papers.nips.cc/paper/9036-modeling-dynamic-functional-connectivity-with-latent-factor-gaussian-processes>
AUTHORS: Lingge Li, Dustin Pluta, Babak Shahbaba, Norbert Fortin, Hernando Ombao, Pierre Baldi
HIGHLIGHT: We present a latent factor Gaussian process model which addresses these challenges by learning a parsimonious representation of connectivity dynamics.
- 742, TITLE: Manipulating a Learning Defender and Ways to Counteract
<https://papers.nips.cc/paper/9037-manipulating-a-learning-defender-and-ways-to-counteract>
AUTHORS: Jiarui Gan, Qingyu Guo, Long Tran-Thanh, Bo An, Michael Wooldridge
HIGHLIGHT: In this paper, we show that, however, these algorithms can be easily manipulated if the attacker responds untruthfully.
- 743, TITLE: Learning-In-The-Loop Optimization: End-To-End Control And Co-Design Of Soft Robots Through Learned Deep Latent Representations
<https://papers.nips.cc/paper/9038-learning-in-the-loop-optimization-end-to-end-control-and-co-design-of-soft-robots-through-learned-deep-latent-representations>
AUTHORS: Andrew Spielberg, Allan Zhao, Yuanming Hu, Tao Du, Wojciech Matusik, Daniela Rus
HIGHLIGHT: We present a learning-in-the-loop co-optimization algorithm in which a latent state representation is learned as the robot figures out how to solve the task.
- 744, TITLE: Learning to Infer Implicit Surfaces without 3D Supervision
<https://papers.nips.cc/paper/9039-learning-to-infer-implicit-surfaces-without-3d-supervision>
AUTHORS: Shichen Liu, Shunsuke Saito, Weikai Chen, Hao Li
HIGHLIGHT: To this end, we propose a novel ray-based field probing technique for efficient image-to-field supervision, as well as a general geometric regularizer for implicit surfaces, which provides natural shape priors in unconstrained regions.
- 745, TITLE: Fast and Accurate Least-Mean-Squares Solvers
<https://papers.nips.cc/paper/9040-fast-and-accurate-least-mean-squares-solvers>
AUTHORS: Ibrahim Jubran, Alaa Maalouf, Dan Feldman
HIGHLIGHT: We suggest an algorithm that gets a finite set of n -dimensional real vectors and returns a weighted subset of $n+1$ vectors whose sum is exactly the same.
- 746, TITLE: Certifiable Robustness to Graph Perturbations
<https://papers.nips.cc/paper/9041-certifiable-robustness-to-graph-perturbations>
AUTHORS: Aleksandar Bojchevski, Stephan G.lemnemann
HIGHLIGHT: We propose the first method for verifying certifiable (non-)robustness to graph perturbations for a general class of models that includes graph neural networks and label/feature propagation.
- 747, TITLE: Fast Convergence of Belief Propagation to Global Optima: Beyond Correlation Decay

- <https://papers.nips.cc/paper/9042-fast-convergence-of-belief-propagation-to-global-optima-beyond-correlation-decay>
AUTHORS: Frederic Koehler
HIGHLIGHT: We show that under a natural initialization, BP converges quickly to the global optimum of the Bethe free energy for Ising models on arbitrary graphs, as long as the Ising model is $\backslash\text{emph}\{\text{ferromagnetic}\}$ (i.e. neighbors prefer to be aligned).
- 748, TITLE: Paradoxes in Fair Machine Learning
<https://papers.nips.cc/paper/9043-paradoxes-in-fair-machine-learning>
AUTHORS: Paul Goelz, Anson Kahng, Ariel D. Procaccia
HIGHLIGHT: We extend equalized odds to the setting of cardinality-constrained fair classification, where we have a bounded amount of a resource to distribute.
- 749, TITLE: Provably Global Convergence of Actor-Critic: A Case for Linear Quadratic Regulator with Ergodic Cost
<https://papers.nips.cc/paper/9044-provably-global-convergence-of-actor-critic-a-case-for-linear-quadratic-regulator-with-ergodic-cost>
AUTHORS: Zhuoran Yang, Yongxin Chen, Mingyi Hong, Zhaoran Wang
HIGHLIGHT: To understand the instability of actor-critic, we focus on its application to linear quadratic regulators, a simple yet fundamental setting of reinforcement learning.
- 750, TITLE: The spiked matrix model with generative priors
<https://papers.nips.cc/paper/9045-the-spiked-matrix-model-with-generative-priors>
AUTHORS: Benjamin Aubin, Bruno Loureiro, Antoine Maillard, Florent Krzakala, Lenka Zdeborov?
HIGHLIGHT: In this paper we study spiked matrix models, where a low-rank matrix is observed through a noisy channel.
- 751, TITLE: Gradient Dynamics of Shallow Univariate ReLU Networks
<https://papers.nips.cc/paper/9046-gradient-dynamics-of-shallow-univariate-relu-networks>
AUTHORS: Francis Williams, Matthew Trager, Daniele Panozzo, Claudio Silva, Denis Zorin, Joan Bruna
HIGHLIGHT: We present a theoretical and empirical study of the gradient dynamics of overparameterized shallow ReLU networks with one-dimensional input, solving least-squares interpolation.
- 752, TITLE: Robust and Communication-Efficient Collaborative Learning
<https://papers.nips.cc/paper/9047-robust-and-communication-efficient-collaborative-learning>
AUTHORS: Amirhossein Reisizadeh, Hossein Taheri, Aryan Mokhtari, Hamed Hassani, Ramtin Pedarsani
HIGHLIGHT: In this paper, we tackle these bottlenecks by proposing a novel decentralized and gradient-based optimization algorithm named as QuanTimed-DSGD.
- 753, TITLE: Multiclass Learning from Contradictions
<https://papers.nips.cc/paper/9048-multiclass-learning-from-contradictions>
AUTHORS: Saurabh Dhar, Vladimir Cherkassky, Mohak Shah
HIGHLIGHT: We introduce the notion of learning from contradictions, a.k.a Universum learning, for multiclass problems and propose a novel formulation for multiclass universum SVM (MU-SVM).
- 754, TITLE: Learning from Trajectories via Subgoal Discovery
<https://papers.nips.cc/paper/9049-learning-from-trajectories-via-subgoal-discovery>
AUTHORS: Sujoy Paul, Jeroen Vanbaer, Amit Roy-Chowdhury
HIGHLIGHT: In this paper, we propose an approach which uses the expert trajectories and learns to decompose the complex main task into smaller sub-goals.
- 755, TITLE: Distributed Low-rank Matrix Factorization With Exact Consensus
<https://papers.nips.cc/paper/9050-distributed-low-rank-matrix-factorization-with-exact-consensus>
AUTHORS: Zhihui Zhu, Qiuwei Li, Xinshuo Yang, Gongguo Tang, Michael B. Wakin
HIGHLIGHT: In this paper, we study low-rank matrix factorization in the distributed setting, where local variables at each node encode parts of the overall matrix factors, and consensus is encouraged among certain such variables.
- 756, TITLE: Online Normalization for Training Neural Networks
<https://papers.nips.cc/paper/9051-online-normalization-for-training-neural-networks>
AUTHORS: Vitaliy Chiley, Ilya Sharapov, Atli Kossou, Urs Koster, Ryan Reece, Sofia Samaniego de la Fuente, Vishal Subbiah, Michael James
HIGHLIGHT: We resolve a theoretical limitation of Batch Normalization by introducing an unbiased technique for computing the gradient of normalized activations.
- 757, TITLE: The Synthesis of XNOR Recurrent Neural Networks with Stochastic Logic

- <https://papers.nips.cc/paper/9052-the-synthesis-of-xnor-recurrent-neural-networks-with-stochastic-logic>
AUTHORS: Arash Ardakani, Zhengyun Ji, Amir Ardakani, Warren Gross
HIGHLIGHT: In this paper, we propose a method that converts all the multiplications in LSTMs to XNOR operations using stochastic computing.
- 758, TITLE: An adaptive Mirror-Prox method for variational inequalities with singular operators
<https://papers.nips.cc/paper/9053-an-adaptive-mirror-prox-method-for-variational-inequalities-with-singular-operators>
AUTHORS: Kimon Antonakopoulos, Veronica Belmega, Panayotis Mertikopoulos
HIGHLIGHT: To address this issue, we propose a novel smoothness condition which we call Bregman smoothness, and which relates the variation of the operator to that of a suitably chosen Bregman function.
- 759, TITLE: N-Gram Graph: Simple Unsupervised Representation for Graphs, with Applications to Molecules
<https://papers.nips.cc/paper/9054-n-gram-graph-simple-unsupervised-representation-for-graphs-with-applications-to-molecules>
AUTHORS: Shengchao Liu, Mehmet F. Demirel, Yingyu Liang
HIGHLIGHT: This paper introduces the N-gram graph, a simple unsupervised representation for molecules.
- 760, TITLE: Characterizing the Exact Behaviors of Temporal Difference Learning Algorithms Using Markov Jump Linear System Theory
<https://papers.nips.cc/paper/9055-characterizing-the-exact-behaviors-of-temporal-difference-learning-algorithms-using-markov-jump-linear-system-theory>
AUTHORS: Bin Hu, Usman Syed
HIGHLIGHT: In this paper, we provide a unified analysis of temporal difference learning algorithms with linear function approximators by exploiting their connections to Markov jump linear systems (MJLS).
- 761, TITLE: Facility Location Problem in Differential Privacy Model Revisited
<https://papers.nips.cc/paper/9056-facility-location-problem-in-differential-privacy-model-revisited>
AUTHORS: Yunus Esencayi, Marco Gaboardi, Shi Li, Di Wang
HIGHLIGHT: In this paper we study the facility location problem in the model of differential privacy (DP) with uniform facility cost.
- 762, TITLE: Energy-Inspired Models: Learning with Sampler-Induced Distributions
<https://papers.nips.cc/paper/9057-energy-inspired-models-learning-with-sampler-induced-distributions>
AUTHORS: John Lawson, George Tucker, Bo Dai, Rajesh Ranganath
HIGHLIGHT: This yields a class of energy-inspired models (EIMs) that incorporate learned energy functions while still providing exact samples and tractable log-likelihood lower bounds.
- 763, TITLE: Finite-time Analysis of Approximate Policy Iteration for the Linear Quadratic Regulator
<https://papers.nips.cc/paper/9058-finite-time-analysis-of-approximate-policy-iteration-for-the-linear-quadratic-regulator>
AUTHORS: Karl Krauth, Stephen Tu, Benjamin Recht
HIGHLIGHT: We study the sample complexity of approximate policy iteration (PI) for the Linear Quadratic Regulator (LQR), building on a recent line of work using LQR as a testbed to understand the limits of reinforcement learning (RL) algorithms on continuous control tasks.
- 764, TITLE: A Universally Optimal Multistage Accelerated Stochastic Gradient Method
<https://papers.nips.cc/paper/9059-a-universally-optimal-multistage-accelerated-stochastic-gradient-method>
AUTHORS: Necdet Serhat Aybat, Alireza Fallah, Mert Gurbuzbalaban, Asuman Ozdaglar
HIGHLIGHT: We propose a novel multistage accelerated algorithm that is universally optimal in the sense that it achieves the optimal rate both in the deterministic and stochastic case and operates without knowledge of noise characteristics.
- 765, TITLE: From deep learning to mechanistic understanding in neuroscience: the structure of retinal prediction
<https://papers.nips.cc/paper/9060-from-deep-learning-to-mechanistic-understanding-in-neuroscience-the-structure-of-retinal-prediction>
AUTHORS: Hidenori Tanaka, Aran Nayebi, Niru Maheswaranathan, Lane McIntosh, Stephen Baccus, Surya Ganguli
HIGHLIGHT: We develop such a systematic approach by combining dimensionality reduction and modern attribution methods for determining the relative importance of interneurons for specific visual computations.
- 766, TITLE: Large Memory Layers with Product Keys
<https://papers.nips.cc/paper/9061-large-memory-layers-with-product-keys>
AUTHORS: Guillaume Lample, Alexandre Sablayrolles, Marc'Aurelio Ranzato, Ludovic Denoyer, Herve Jegou
HIGHLIGHT: This paper introduces a structured memory which can be easily integrated into a neural network.

- 767, TITLE: Learning Deterministic Weighted Automata with Queries and Counterexamples
<https://papers.nips.cc/paper/9062-learning-deterministic-weighted-automata-with-queries-and-counterexamples>
AUTHORS: Gail Weiss, Yoav Goldberg, Eran Yahav
HIGHLIGHT: We present an algorithm for reconstruction of a probabilistic deterministic finite automaton (PDFA) from a given black-box language model, such as a recurrent neural network (RNN).
- 768, TITLE: Wide Neural Networks of Any Depth Evolve as Linear Models Under Gradient Descent
<https://papers.nips.cc/paper/9063-wide-neural-networks-of-any-depth-evolve-as-linear-models-under-gradient-descent>
AUTHORS: Jaehoon Lee, Lechao Xiao, Samuel Schoenholz, Yasaman Bahri, Roman Novak, Jascha Sohl-Dickstein, Jeffrey Pennington
HIGHLIGHT: In this work, we show that for wide neural networks the learning dynamics simplify considerably and that, in the infinite width limit, they are governed by a linear model obtained from the first-order Taylor expansion of the network around its initial parameters.
- 769, TITLE: Time/Accuracy Tradeoffs for Learning a ReLU with respect to Gaussian Marginals
<https://papers.nips.cc/paper/9064-timeaccuracy-tradeoffs-for-learning-a-relu-with-respect-to-gaussian-marginals>
AUTHORS: Surbhi Goel, Sushrut Karmalkar, Adam Klivans
HIGHLIGHT: We consider the problem of computing the best-fitting ReLU with respect to square-loss on a training set when the examples have been drawn according to a spherical Gaussian distribution (the labels can be arbitrary).
- 770, TITLE: Visualizing and Measuring the Geometry of BERT
<https://papers.nips.cc/paper/9065-visualizing-and-measuring-the-geometry-of-bert>
AUTHORS: Emily Reif, Ann Yuan, Martin Wattenberg, Fernanda B. Viegas, Andy Coenen, Adam Pearce, Been Kim
HIGHLIGHT: This paper describes qualitative and quantitative investigations of one particularly effective model, BERT.
- 771, TITLE: Self-Critical Reasoning for Robust Visual Question Answering
<https://papers.nips.cc/paper/9066-self-critical-reasoning-for-robust-visual-question-answering>
AUTHORS: Jialin Wu, Raymond Mooney
HIGHLIGHT: To address this issue, we introduce a self-critical training objective that ensures that visual explanations of correct answers match the most influential image regions more than other competitive answer candidates.
- 772, TITLE: Learning to Screen
<https://papers.nips.cc/paper/9067-learning-to-screen>
AUTHORS: Alon Cohen, Avinatan Hassidim, Haim Kaplan, Yishay Mansour, Shay Moran
HIGHLIGHT: We model such scenarios as an assignment problem between items (candidates) and categories (departments): the items arrive one-by-one in an online manner, and upon processing each item the algorithm decides, based on its value and the categories it can be matched with, whether to retain or discard it (this decision is irrevocable).
- 773, TITLE: A Communication Efficient Stochastic Multi-Block Alternating Direction Method of Multipliers
<https://papers.nips.cc/paper/9068-a-communication-efficient-stochastic-multi-block-alternating-direction-method-of-multipliers>
AUTHORS: Hao Yu
HIGHLIGHT: In this paper, we propose a new parallel multi-block stochastic ADMM for distributed stochastic optimization, where each node is only required to perform simple stochastic gradient descent updates.
- 774, TITLE: A Little Is Enough: Circumventing Defenses For Distributed Learning
<https://papers.nips.cc/paper/9069-a-little-is-enough-circumventing-defenses-for-distributed-learning>
AUTHORS: Gilad Baruch, Moran Baruch, Yoav Goldberg
HIGHLIGHT: We observe that if the empirical variance between the gradients of workers is high enough, an attacker could take advantage of this and launch a non-omniscient attack that operates within the population variance.
- 775, TITLE: Error Correcting Output Codes Improve Probability Estimation and Adversarial Robustness of Deep Neural Networks
<https://papers.nips.cc/paper/9070-error-correcting-output-codes-improve-probability-estimation-and-adversarial-robustness-of-deep-neural-networks>
AUTHORS: Gunjan Verma, Ananthram Swami
HIGHLIGHT: In this paper, we propose a fundamentally different approach which instead changes the way the output is represented and decoded.
- 776, TITLE: A Robust Non-Clairvoyant Dynamic Mechanism for Contextual Auctions
<https://papers.nips.cc/paper/9071-a-robust-non-clairvoyant-dynamic-mechanism-for-contextual-auctions>

- AUTHORS: Yuan Deng, S?bastien Lahaie, Vahab Mirrokni
HIGHLIGHT: In this paper, we consider the problem of contextual auctions where the seller gradually learns a model of the buyer's valuation as a function of the context (e.g., item features) and seeks a pricing policy that optimizes revenue.
- 777, TITLE: Finite-Sample Analysis for SARSA with Linear Function Approximation
<https://papers.nips.cc/paper/9072-finite-sample-analysis-for-sarsa-with-linear-function-approximation>
AUTHORS: Shaofeng Zou, Tengyu Xu, Yingbin Liang
HIGHLIGHT: In this paper, we develop a novel technique to explicitly characterize the stochastic bias of a type of stochastic approximation procedures with time-varying Markov transition kernels.
- 778, TITLE: Who is Afraid of Big Bad Minima? Analysis of gradient-flow in spiked matrix-tensor models
<https://papers.nips.cc/paper/9073-who-is-afraid-of-big-bad-minima-analysis-of-gradient-flow-in-spiked-matrix-tensor-models>
AUTHORS: Stefano Sarao Mannelli, Giulio Biroli, Chiara Cammarota, Florent Krzakala, Lenka Zdeborov?
HIGHLIGHT: Here we present a quantitative theory explaining this behaviour in a spiked matrix-tensor model.
- 779, TITLE: Graph Structured Prediction Energy Networks
<https://papers.nips.cc/paper/9074-graph-structured-prediction-energy-networks>
AUTHORS: Colin Graber, Alexander Schwing
HIGHLIGHT: To address this shortcoming, we introduce 'Graph Structured Prediction Energy Networks,' for which we develop inference techniques that allow to both model explicit local and implicit higher-order correlations while maintaining tractability of inference.
- 780, TITLE: Private Learning Implies Online Learning: An Efficient Reduction
<https://papers.nips.cc/paper/9075-private-learning-implies-online-learning-an-efficient-reduction>
AUTHORS: Alon Gonen, Elad Hazan, Shay Moran
HIGHLIGHT: In this paper we resolve this open question in the context of pure differential privacy.
- 781, TITLE: Graph Agreement Models for Semi-Supervised Learning
<https://papers.nips.cc/paper/9076-graph-agreement-models-for-semi-supervised-learning>
AUTHORS: Otilia Stretcu, Krishnamurthy Viswanathan, Dana Movshovitz-Attias, Emmanouil Platanios, Sujith Ravi, Andrew Tomkins
HIGHLIGHT: To address this, we propose Graph Agreement Models (GAM), which introduces an auxiliary model that predicts the probability of two nodes sharing the same label as a learned function of their features.
- 782, TITLE: Latent distance estimation for random geometric graphs
<https://papers.nips.cc/paper/9077-latent-distance-estimation-for-random-geometric-graphs>
AUTHORS: Ernesto Araya Valdivia, De Castro Yohann
HIGHLIGHT: We introduce a spectral estimator of the pairwise distance between latent points and we prove that its rate of convergence is the same as the nonparametric estimation of a function on \mathbb{S}^{d-1} , up to a logarithmic factor.
- 783, TITLE: Seeing the Wind: Visual Wind Speed Prediction with a Coupled Convolutional and Recurrent Neural Network
<https://papers.nips.cc/paper/9078-seeing-the-wind-visual-wind-speed-prediction-with-a-coupled-convolutional-and-recurrent-neural-network>
AUTHORS: Jennifer Cardona, Michael Howland, John Dabiri
HIGHLIGHT: Here, we demonstrate a coupled convolutional neural network and recurrent neural network architecture that extracts the wind speed encoded in visually recorded flow-structure interactions of a flag and tree in naturally occurring wind.
- 784, TITLE: The Functional Neural Process
<https://papers.nips.cc/paper/9079-the-functional-neural-process>
AUTHORS: Christos Louizos, Xiahn Shi, Klamer Schutte, Max Welling
HIGHLIGHT: We present a new family of exchangeable stochastic processes, the Functional Neural Processes (FNPs).
- 785, TITLE: Recurrent Registration Neural Networks for Deformable Image Registration
<https://papers.nips.cc/paper/9080-recurrent-registration-neural-networks-for-deformable-image-registration>
AUTHORS: Robin Sandk?hler, Simon Andermatt, Grzegorz Bauman, Sylvia Nyilas, Christoph Jud, Philippe C. Cattin
HIGHLIGHT: We reformulate the pairwise registration problem as a recursive sequence of successive alignments.
- 786, TITLE: Unsupervised State Representation Learning in Atari
<https://papers.nips.cc/paper/9081-unsupervised-state-representation-learning-in-atari>
AUTHORS: Ankesh Anand, Evan Racah, Sherjil Ozair, Yoshua Bengio, Marc-Alexandre C?t?, R Devon Hjelm

HIGHLIGHT: We introduce a method that tries to learn better state representations by maximizing mutual information across spatially and temporally distinct features of a neural encoder of the observations.

787, **TITLE:** Unlocking Fairness: a Trade-off Revisited
<https://papers.nips.cc/paper/9082-unlocking-fairness-a-trade-off-revisited>

AUTHORS: Michael Wick, swetasudha panda, Jean-Baptiste Tristan

HIGHLIGHT: We investigate fairness and accuracy, but this time under a variety of controlled conditions in which we vary the amount and type of bias.

788, **TITLE:** Fisher Efficient Inference of Intractable Models
<https://papers.nips.cc/paper/9083-fisher-efficient-inference-of-intractable-models>

AUTHORS: Song Liu, Takafumi Kanamori, Wittawat Jitkrittum, Yu Chen

HIGHLIGHT: In this paper, we derive a Discriminative Likelihood Estimator (DLE) from the Kullback-Leibler divergence minimization criterion implemented via density ratio estimation and a Stein operator.

789, **TITLE:** Thompson Sampling and Approximate Inference
<https://papers.nips.cc/paper/9084-thompson-sampling-and-approximate-inference>

AUTHORS: My Phan, Yasin Abbasi Yadkori, Justin Domke

HIGHLIGHT: We study the effects of approximate inference on the performance of Thompson sampling in the k -armed bandit problems.

790, **TITLE:** PRNet: Self-Supervised Learning for Partial-to-Partial Registration

<https://papers.nips.cc/paper/9085-prnet-self-supervised-learning-for-partial-to-partial-registration>

AUTHORS: Yue Wang, Justin M. Solomon

HIGHLIGHT: We present a simple, flexible, and general framework titled Partial Registration Network (PRNet), for partial-to-partial point cloud registration.

791, **TITLE:** Surrogate Objectives for Batch Policy Optimization in One-step Decision Making

<https://papers.nips.cc/paper/9086-surrogate-objectives-for-batch-policy-optimization-in-one-step-decision-making>

AUTHORS: Minmin Chen, Ramki Gummadi, Chris Harris, Dale Schuurmans

HIGHLIGHT: We investigate batch policy optimization for cost-sensitive classification and contextual bandits---two related tasks that obviate exploration but require generalizing from observed rewards to action selections in unseen contexts.

792, **TITLE:** Modelling heterogeneous distributions with an Uncountable Mixture of Asymmetric Laplacians

<https://papers.nips.cc/paper/9087-modelling-heterogeneous-distributions-with-an-uncountable-mixture-of-asymmetric-laplacians>

AUTHORS: Axel Brando, Jose A. Rodriguez, Jordi Vitria, Alberto Rubio Muoz

HIGHLIGHT: In this paper, we propose a generic deep learning framework that learns an Uncountable Mixture of Asymmetric Laplacians (UMAL), which will allow us to estimate heterogeneous distributions of the output variable and shows its connections to quantile regression.

793, **TITLE:** Learning Macroscopic Brain Connectomes via Group-Sparse Factorization

<https://papers.nips.cc/paper/9088-learning-macroscopic-brain-connectomes-via-group-sparse-factorization>

AUTHORS: Farzane Aminmansour, Andrew Patterson, Lei Le, Yisu Peng, Daniel Mitchell, Franco Pestilli, Cesar F. Caiafa, Russell Greiner, Martha White

HIGHLIGHT: In this work, we explore a framework that facilitates applying learning algorithms to automatically extract brain connectomes.

794, **TITLE:** Approximating the Permanent by Sampling from Adaptive Partitions

<https://papers.nips.cc/paper/9089-approximating-the-permanent-by-sampling-from-adaptive-partitions>

AUTHORS: Jonathan Kuck, Tri Dao, Hamid Rezaatfighi, Ashish Sabharwal, Stefano Ermon

HIGHLIGHT: We present ADAPART, a simple and efficient method for exact sampling of permutations, each associated with a weight as determined by a matrix.

795, **TITLE:** Retrosynthesis Prediction with Conditional Graph Logic Network

<https://papers.nips.cc/paper/9090-retrosynthesis-prediction-with-conditional-graph-logic-network>

AUTHORS: Hanjun Dai, Chengtao Li, Connor Coley, Bo Dai, Le Song

HIGHLIGHT: In this work, we propose a new approach to this task using the Conditional Graph Logic Network, a conditional graphical model built upon graph neural networks that learns when rules from reaction templates should be applied, implicitly considering whether the resulting reaction would be both chemically feasible and strategic.

- 796, TITLE: Procrastinating with Confidence: Near-Optimal, Anytime, Adaptive Algorithm Configuration
<https://papers.nips.cc/paper/9091-procrastinating-with-confidence-near-optimal-anytime-adaptive-algorithm-configuration>
AUTHORS: Robert Kleinberg, Kevin Leyton-Brown, Brendan Lucier, Devon Graham
HIGHLIGHT: This paper introduces a new algorithm, "Structured Procrastination with Confidence", that preserves the near-optimality and anytime properties of Structured Procrastination while adding adaptivity.
- 797, TITLE: Online Learning via the Differential Privacy Lens
<https://papers.nips.cc/paper/9092-online-learning-via-the-differential-privacy-lens>
AUTHORS: Jacob D. Abernethy, Young Hun Jung, Chansoo Lee, Audra McMillan, Ambuj Tewari
HIGHLIGHT: In this paper, we use differential privacy as a lens to examine online learning in both full and partial information settings.
- 798, TITLE: PerspectiveNet: 3D Object Detection from a Single RGB Image via Perspective Points
<https://papers.nips.cc/paper/9093-perspectivenet-3d-object-detection-from-a-single-rgb-image-via-perspective-points>
AUTHORS: Siyuan Huang, Yixin Chen, Tao Yuan, Siyuan Qi, Yixin Zhu, Song-Chun Zhu
HIGHLIGHT: To address this challenge, we propose to adopt perspective points as a new intermediate representation for 3D object detection, defined as the 2D projections of local Manhattan 3D keypoints to locate an object; these perspective points satisfy geometric constraints imposed by the perspective projection.
- 799, TITLE: Parameter elimination in particle Gibbs sampling
<https://papers.nips.cc/paper/9094-parameter-elimination-in-particle-gibbs-sampling>
AUTHORS: Anna Wigren, Riccardo Sven Risuleo, Lawrence Murray, Fredrik Lindsten
HIGHLIGHT: We focus on particle Gibbs (PG) and particle Gibbs with ancestor sampling (PGAS), improving their performance beyond that of the ideal Gibbs sampler (which they approximate) by marginalizing out one or more parameters.
- 800, TITLE: This Looks Like That: Deep Learning for Interpretable Image Recognition
<https://papers.nips.cc/paper/9095-this-looks-like-that-deep-learning-for-interpretable-image-recognition>
AUTHORS: Chaofan Chen, Oscar Li, Daniel Tao, Alina Barnett, Cynthia Rudin, Jonathan K. Su
HIGHLIGHT: In this work, we introduce a deep network architecture -- prototypical part network (ProtoPNet), that reasons in a similar way: the network dissects the image by finding prototypical parts, and combines evidence from the prototypes to make a final classification.
- 801, TITLE: Adaptively Aligned Image Captioning via Adaptive Attention Time
<https://papers.nips.cc/paper/9096-adaptively-aligned-image-captioning-via-adaptive-attention-time>
AUTHORS: Lun Huang, Wenmin Wang, Yaxian Xia, Jie Chen
HIGHLIGHT: In this paper, we propose a novel attention model, namely Adaptive Attention Time (AAT), to align the source and the target adaptively for image captioning.
- 802, TITLE: Accurate Uncertainty Estimation and Decomposition in Ensemble Learning
<https://papers.nips.cc/paper/9097-accurate-uncertainty-estimation-and-decomposition-in-ensemble-learning>
AUTHORS: Jeremiah Liu, John Paisley, Marianthi-Anna Kioumourtzoglou, Brent Coull
HIGHLIGHT: We introduce a Bayesian nonparametric ensemble (BNE) approach that augments an existing ensemble model to account for different sources of model uncertainty.
- 803, TITLE: Learning Bayesian Networks with Low Rank Conditional Probability Tables
<https://papers.nips.cc/paper/9098-learning-bayesian-networks-with-low-rank-conditional-probability-tables>
AUTHORS: Adarsh Barik, Jean Honorio
HIGHLIGHT: In this paper, we provide a method to learn the directed structure of a Bayesian network using data.
- 804, TITLE: Equal Opportunity in Online Classification with Partial Feedback
<https://papers.nips.cc/paper/9099-equal-opportunity-in-online-classification-with-partial-feedback>
AUTHORS: Yahav Bechavod, Katrina Ligett, Aaron Roth, Bo Waggoner, Steven Z. Wu
HIGHLIGHT: We study an online classification problem with partial feedback in which individuals arrive one at a time from a fixed but unknown distribution, and must be classified as positive or negative.
- 805, TITLE: Modeling Expectation Violation in Intuitive Physics with Coarse Probabilistic Object Representations
<https://papers.nips.cc/paper/9100-modeling-expectation-violation-in-intuitive-physics-with-coarse-probabilistic-object-representations>
AUTHORS: Kevin Smith, Lingjie Mei, Shunyu Yao, Jiajun Wu, Elizabeth Spelke, Josh Tenenbaum, Tomer Ullman
HIGHLIGHT: We propose ADEPT, a model that uses a coarse (approximate geometry) object-centric representation for dynamic 3D scene understanding.

806, TITLE: Neural Multisensory Scene Inference
<https://papers.nips.cc/paper/9101-neural-multisensory-scene-inference>
AUTHORS: Jae Hyun Lim, Pedro O. O. Pinheiro, Negar Rostamzadeh, Chris Pal, Sungjin Ahn
HIGHLIGHT: In this paper, we propose the Generative Multisensory Network (GMN) for learning latent representations of 3D scenes which are partially observable through multiple sensory modalities.

807, TITLE: Regret Bounds for Thompson Sampling in Episodic Restless Bandit Problems
<https://papers.nips.cc/paper/9102-regret-bounds-for-thompson-sampling-in-episodic-restless-bandit-problems>
AUTHORS: Young Hun Jung, Ambuj Tewari
HIGHLIGHT: In this paper, we analyze the performance of Thompson sampling in episodic restless bandits with unknown parameters.

808, TITLE: What Can ResNet Learn Efficiently, Going Beyond Kernels?
<https://papers.nips.cc/paper/9103-what-can-resnet-learn-efficiently-going-beyond-kernels>
AUTHORS: Zeyuan Allen-Zhu, Yuanzhi Li
HIGHLIGHT: We prove neural networks can efficiently learn a notable class of functions, including those defined by three-layer residual networks with smooth activations, without any distributional assumption.

809, TITLE: Better Transfer Learning with Inferred Successor Maps
<https://papers.nips.cc/paper/9104-better-transfer-learning-with-inferred-successor-maps>
AUTHORS: Tamas Madarasz, Tim Behrens
HIGHLIGHT: We thus provide a novel algorithmic approach for multi-task learning, as well as a common normative framework that links together these different characteristics of the brain's spatial representation.

810, TITLE: Unsupervised Co-Learning on G -Manifolds Across Irreducible Representations
<https://papers.nips.cc/paper/9105-unsupervised-co-learning-on-g-manifolds-across-irreducible-representations>
AUTHORS: Yifeng Fan, Tingran Gao, Zhizhen Jane Zhao
HIGHLIGHT: We introduce a novel co-learning paradigm for manifolds naturally admitting an action of a transformation group G , motivated by recent developments on learning a manifold from attached fibre bundle structures.

811, TITLE: Defending Against Neural Fake News
<https://papers.nips.cc/paper/9106-defending-against-neural-fake-news>
AUTHORS: Rowan Zellers, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, Yejin Choi
HIGHLIGHT: We thus present a model for controllable text generation called Grover.

812, TITLE: Sample Adaptive MCMC
<https://papers.nips.cc/paper/9107-sample-adaptive-mcmc>
AUTHORS: Michael Zhu
HIGHLIGHT: In this paper, we present Sample Adaptive MCMC (SA-MCMC), a MCMC method based on a reversible Markov chain for $\pi^{\otimes N}$ that uses an adaptive proposal distribution based on the current state of N points and a sequential substitution procedure with one new likelihood evaluation per iteration and at most one updated point each iteration.

813, TITLE: A Stochastic Composite Gradient Method with Incremental Variance Reduction
<https://papers.nips.cc/paper/9108-a-stochastic-composite-gradient-method-with-incremental-variance-reduction>
AUTHORS: Junyu Zhang, Lin Xiao
HIGHLIGHT: We propose a stochastic composite gradient method that employs incremental variance-reduced estimators for both the inner vector mapping and its Jacobian.

814, TITLE: Nonparametric Density Estimation & Convergence Rates for GANs under Besov IPM Losses
<https://papers.nips.cc/paper/9109-nonparametric-density-estimation-convergence-rates-for-gans-under-besov-ipm-losses>
AUTHORS: Ananya Uppal, Shashank Singh, Barnabas Poczos
HIGHLIGHT: We study the problem of estimating a nonparametric probability distribution under a family of losses called Besov IPMs.

815, TITLE: STAR-Caps: Capsule Networks with Straight-Through Attentive Routing
<https://papers.nips.cc/paper/9110-star-caps-capsule-networks-with-straight-through-attentive-routing>
AUTHORS: Karim Ahmed, Lorenzo Torresani
HIGHLIGHT: In this work, we propose Star-Caps a capsule-based network that exploits a straight-through attentive routing to address the drawbacks of capsule networks.

- 816, TITLE: Limitations of Lazy Training of Two-layers Neural Network
<https://papers.nips.cc/paper/9111-limitations-of-lazy-training-of-two-layers-neural-network>
AUTHORS: Song Mei, Theodor Misiakiewicz, Behrooz Ghorbani, Andrea Montanari
HIGHLIGHT: We study the supervised learning problem under either of the following two models: (1) Feature vectors x_i are d -dimensional Gaussian and responses are $y_i = f_*(x_i)$ for f_* an unknown quadratic function; (2) Feature vectors x_i are distributed as a mixture of two d -dimensional centered Gaussians, and y_i 's are the corresponding class labels.
- 817, TITLE: Reconciling meta-learning and continual learning with online mixtures of tasks
<https://papers.nips.cc/paper/9112-reconciling-meta-learning-and-continual-learning-with-online-mixtures-of-tasks>
AUTHORS: Ghassen Jerfel, Erin Grant, Tom Griffiths, Katherine A. Heller
HIGHLIGHT: We use the connection between gradient-based meta-learning and hierarchical Bayes to propose a Dirichlet process mixture of hierarchical Bayesian models over the parameters of an arbitrary parametric model such as a neural network.
- 818, TITLE: Distributionally Robust Optimization and Generalization in Kernel Methods
<https://papers.nips.cc/paper/9113-distributionally-robust-optimization-and-generalization-in-kernel-methods>
AUTHORS: Matthew Staib, Stefanie Jegelka
HIGHLIGHT: In this paper, we study DRO with uncertainty sets measured via maximum mean discrepancy (MMD).
- 819, TITLE: A General Theory of Equivariant CNNs on Homogeneous Spaces
<https://papers.nips.cc/paper/9114-a-general-theory-of-equivariant-cnns-on-homogeneous-spaces>
AUTHORS: Taco S. Cohen, Mario Geiger, Maurice Weiler
HIGHLIGHT: We present a general theory of Group equivariant Convolutional Neural Networks (G-CNNs) on homogeneous spaces such as Euclidean space and the sphere.
- 820, TITLE: Trivializations for Gradient-Based Optimization on Manifolds
<https://papers.nips.cc/paper/9115-trivializations-for-gradient-based-optimization-on-manifolds>
AUTHORS: Mario Lezcano Casado
HIGHLIGHT: We introduce a framework to study the transformation of problems with manifold constraints into unconstrained problems through parametrizations in terms of a Euclidean space.
- 821, TITLE: Write, Execute, Assess: Program Synthesis with a REPL
<https://papers.nips.cc/paper/9116-write-execute-assess-program-synthesis-with-a-repl>
AUTHORS: Kevin Ellis, Maxwell Nye, Yewen Pu, Felix Sosa, Josh Tenenbaum, Armando Solar-Lezama
HIGHLIGHT: We present a neural program synthesis approach integrating components which write, execute, and assess code to navigate the search space of possible programs.
- 822, TITLE: A Meta-Analysis of Overfitting in Machine Learning
<https://papers.nips.cc/paper/9117-a-meta-analysis-of-overfitting-in-machine-learning>
AUTHORS: Rebecca Roelofs, Vaishal Shankar, Benjamin Recht, Sara Fridovich-Keil, Moritz Hardt, John Miller, Ludwig Schmidt
HIGHLIGHT: We conduct the first large meta-analysis of overfitting due to test set reuse in the machine learning community.
- 823, TITLE: (Nearly) Efficient Algorithms for the Graph Matching Problem on Correlated Random Graphs
<https://papers.nips.cc/paper/9118-nearly-efficient-algorithms-for-the-graph-matching-problem-on-correlated-random-graphs>
AUTHORS: Boaz Barak, Chi-Ning Chou, Zhixian Lei, Tselil Schramm, Yueqi Sheng
HIGHLIGHT: We consider the graph matching/similarity problem of determining how similar two given graphs G_0, G_1 are and recovering the permutation π on the vertices of G_1 that minimizes the symmetric difference between the edges of G_0 and $\pi(G_1)$.
- 824, TITLE: Preference-Based Batch and Sequential Teaching: Towards a Unified View of Models
<https://papers.nips.cc/paper/9119-preference-based-batch-and-sequential-teaching-towards-a-unified-view-of-models>
AUTHORS: Farnam Mansouri, Yuxin Chen, Ara Vartanian, Jerry Zhu, Adish Singla
HIGHLIGHT: To better understand the connections between these different batch and sequential models, we develop a novel framework which captures the teaching process via preference functions
- 825, TITLE: Online Continuous Submodular Maximization: From Full-Information to Bandit Feedback
<https://papers.nips.cc/paper/9120-online-continuous-submodular-maximization-from-full-information-to-bandit-feedback>
AUTHORS: Mingrui Zhang, Lin Chen, Hamed Hassani, Amin Karbasi
HIGHLIGHT: In this paper, we propose three online algorithms for submodular maximization.

826, TITLE: Sampling Networks and Aggregate Simulation for Online POMDP Planning
<https://papers.nips.cc/paper/9121-sampling-networks-and-aggregate-simulation-for-online-pomdp-planning>
AUTHORS: Hao(Jackson) Cui, Roni Khadon
HIGHLIGHT: The paper introduces a new algorithm for planning in partially observable Markov decision processes (POMDP) based on the idea of aggregate simulation.

827, TITLE: Correlation in Extensive-Form Games: Saddle-Point Formulation and Benchmarks
<https://papers.nips.cc/paper/9122-correlation-in-extensive-form-games-saddle-point-formulation-and-benchmarks>
AUTHORS: Gabriele Farina, Chun Kai Ling, Fei Fang, Tuomas Sandholm
HIGHLIGHT: To showcase how this novel formulation can inspire new algorithms to compute EFCEs, we propose a simple subgradient descent method which exploits this formulation and structural properties of EFCEs.

828, TITLE: GNNExplainer: Generating Explanations for Graph Neural Networks
<https://papers.nips.cc/paper/9123-gnnexplainer-generating-explanations-for-graph-neural-networks>
AUTHORS: Zhitao Ying, Dylan Bourgeois, Jiaxuan You, Marinka Zitnik, Jure Leskovec
HIGHLIGHT: Here we propose GNNExplainer, the first general, model-agnostic approach for providing interpretable explanations for predictions of any GNN-based model on any graph-based machine learning task.

829, TITLE: Linear Stochastic Bandits Under Safety Constraints
<https://papers.nips.cc/paper/9124-linear-stochastic-bandits-under-safety-constraints>
AUTHORS: Sanae Amani, Mahnoosh Alizadeh, Christos Thrampoulidis
HIGHLIGHT: In this paper, we formulate a linear stochastic multi-armed bandit problem with safety constraints that depend (linearly) on an unknown parameter vector.

830, TITLE: A coupled autoencoder approach for multi-modal analysis of cell types
<https://papers.nips.cc/paper/9125-a-coupled-autoencoder-approach-for-multi-modal-analysis-of-cell-types>
AUTHORS: Rohan Gala, Nathan Gouwens, Zizhen Yao, Agata Budzillo, Osnat Penn, Bosiljka Tasic, Gabe Murphy, Hongkui Zeng, Uygur S?mb?l
HIGHLIGHT: We pose this issue of cross-modal alignment as an optimization problem and develop an approach based on coupled training of autoencoders as a framework for such analyses.

831, TITLE: Towards Automatic Concept-based Explanations
<https://papers.nips.cc/paper/9126-towards-automatic-concept-based-explanations>
AUTHORS: Amirata Ghorbani, James Wexler, James Y. Zou, Been Kim
HIGHLIGHT: In this work, we propose principles and desiderata for \emph{concept} based explanation, which goes beyond per-sample features to identify higher level human-understandable concepts that apply across the entire dataset.

832, TITLE: Deep Generative Video Compression
<https://papers.nips.cc/paper/9127-deep-generative-video-compression>
AUTHORS: Salvator Lombardo, JUN HAN, Christopher Schroers, Stephan Mandt
HIGHLIGHT: Here, we propose an end-to-end, deep generative modeling approach to compress temporal sequences with a focus on video.

833, TITLE: Budgeted Reinforcement Learning in Continuous State Space
<https://papers.nips.cc/paper/9128-budgeted-reinforcement-learning-in-continuous-state-space>
AUTHORS: Nicolas Carrara, Edouard Leurent, Romain Laroche, Tanguy Urvoy, Odalric-Ambrym Maillard, Olivier Pietquin
HIGHLIGHT: This observation allows us to introduce natural extensions of Deep Reinforcement Learning algorithms to address large-scale BMDPs.

834, TITLE: Discovery of Useful Questions as Auxiliary Tasks
<https://papers.nips.cc/paper/9129-discovery-of-useful-questions-as-auxiliary-tasks>
AUTHORS: Vivek Veeriah, Matteo Hessel, Zhongwen Xu, Janarthanan Rajendran, Richard L. Lewis, Junhyuk Oh, Hado P. van Hasselt, David Silver, Satinder Singh
HIGHLIGHT: We present a novel method for a reinforcement learning (RL) agent to discover questions formulated as general value functions or GVFs, a fairly rich form of knowledge representation.

835, TITLE: Sinkhorn Barycenters with Free Support via Frank-Wolfe Algorithm
<https://papers.nips.cc/paper/9130-sinkhorn-barycenters-with-free-support-via-frank-wolfe-algorithm>
AUTHORS: Giulia Luise, Saverio Salzo, Massimiliano Pontil, Carlo Ciliberto

HIGHLIGHT: We present a novel algorithm to estimate the barycenter of arbitrary probability distributions with respect to the Sinkhorn divergence.

836, **TITLE:** Finding the Needle in the Haystack with Convolutions: on the benefits of architectural bias
<https://papers.nips.cc/paper/9131-finding-the-needle-in-the-haystack-with-convolutions-on-the-benefits-of-architectural-bias>
AUTHORS: Stéphane d'Ascoli, Levent Sagun, Giulio Biroli, Joan Bruna
HIGHLIGHT: The aim of this work is to understand this fact through the lens of dynamics in the loss landscape.

837, **TITLE:** Correlation clustering with local objectives
<https://papers.nips.cc/paper/9132-correlation-clustering-with-local-objectives>
AUTHORS: Sanchit Kalhan, Konstantin Makarychev, Timothy Zhou
HIGHLIGHT: In this paper, we study algorithms for minimizing ℓ_1 norms ($q \geq 1$) of the disagreements vector for both arbitrary and complete graphs.

838, **TITLE:** Multiclass Performance Metric Elicitation
<https://papers.nips.cc/paper/9133-multiclass-performance-metric-elicitation>
AUTHORS: Gaurush Hiranandani, Shant Boodaghians, Ruta Mehta, Oluwasanmi O. Koyejo
HIGHLIGHT: In this paper, we propose novel strategies for eliciting multiclass classification performance metrics using only relative preference feedback.

839, **TITLE:** Algorithmic Analysis and Statistical Estimation of SLOPE via Approximate Message Passing
<https://papers.nips.cc/paper/9134-algorithmic-analysis-and-statistical-estimation-of-slope-via-approximate-message-passing>
AUTHORS: Zhiqi Bu, Jason Klusowski, Cynthia Rush, Weijie Su
HIGHLIGHT: In this paper, we develop an asymptotically exact characterization of the SLOPE solution under Gaussian random designs through solving the SLOPE problem using approximate message passing (AMP).

840, **TITLE:** Explicit Explore-Exploit Algorithms in Continuous State Spaces
<https://papers.nips.cc/paper/9135-explicit-explore-exploit-algorithms-in-continuous-state-spaces>
AUTHORS: Mikael Henaff
HIGHLIGHT: We present a new model-based algorithm for reinforcement learning (RL) which consists of explicit exploration and exploitation phases, and is applicable in large or infinite state spaces.

841, **TITLE:** ADDIS: an adaptive discarding algorithm for online FDR control with conservative nulls
<https://papers.nips.cc/paper/9136-addis-an-adaptive-discarding-algorithm-for-online-fdr-control-with-conservative-nulls>
AUTHORS: Jinjin Tian, Aaditya Ramdas
HIGHLIGHT: In this work, we introduce a new adaptive discarding method called ADDIS that provably controls the FDR and achieves the best of both worlds: it enjoys appreciable power increase over all existing methods if nulls are conservative (the practical case), and rarely loses power if nulls are exactly uniformly distributed (the ideal case).

842, **TITLE:** Slice-based Learning: A Programming Model for Residual Learning in Critical Data Slices
<https://papers.nips.cc/paper/9137-slice-based-learning-a-programming-model-for-residual-learning-in-critical-data-slices>
AUTHORS: Vincent Chen, Sen Wu, Alexander J. Ratner, Jen Weng, Christopher R?
HIGHLIGHT: We propose Slice-based Learning, a new programming model in which the slicing function (SF), a programmer abstraction, is used to specify additional model capacity for each slice.

843, **TITLE:** Don't Blame the ELBO! A Linear VAE Perspective on Posterior Collapse
<https://papers.nips.cc/paper/9138-dont-blame-the-elbo-a-linear-vae-perspective-on-posterior-collapse>
AUTHORS: James Lucas, George Tucker, Roger B. Grosse, Mohammad Norouzi
HIGHLIGHT: This paper presents a simple and intuitive explanation for posterior collapse through the analysis of linear VAEs and their direct correspondence with Probabilistic PCA (pPCA).

844, **TITLE:** Language as an Abstraction for Hierarchical Deep Reinforcement Learning
<https://papers.nips.cc/paper/9139-language-as-an-abstraction-for-hierarchical-deep-reinforcement-learning>
AUTHORS: YiDing Jiang, Shixiang (Shane) Gu, Kevin P. Murphy, Chelsea Finn
HIGHLIGHT: In this paper, we propose to use language as the abstraction, as it provides unique compositional structure, enabling fast learning and combinatorial generalization, while retaining tremendous flexibility, making it suitable for a variety of problems.

845, **TITLE:** Efficient online learning with kernels for adversarial large scale problems
<https://papers.nips.cc/paper/9140-efficient-online-learning-with-kernels-for-adversarial-large-scale-problems>

AUTHORS: R?mi J?z?quel, Pierre Gaillard, Alessandro Rudi
HIGHLIGHT: Our contributions are twofold: 1) For the Gaussian kernel, we propose to build the basis beforehand (independently of the data) through Taylor expansion.

846, TITLE: A Linearly Convergent Method for Non-Smooth Non-Convex Optimization on the Grassmannian with Applications to Robust Subspace and Dictionary Learning
<https://papers.nips.cc/paper/9141-a-linearly-convergent-method-for-non-smooth-non-convex-optimization-on-the-grassmannian-with-applications-to-robust-subspace-and-dictionary-learning>
AUTHORS: Zhihui Zhu, Tianyu Ding, Daniel Robinson, Manolis Tsakiris, Ren? Vidal
HIGHLIGHT: In this paper we show that if the objective satisfies a certain Riemannian regularity condition with respect to some point in the Grassmannian, then a Riemannian subgradient method with appropriate initialization and geometrically diminishing step size converges at a linear rate to that point.

847, TITLE: ObjectNet: A large-scale bias-controlled dataset for pushing the limits of object recognition models
<https://papers.nips.cc/paper/9142-objectnet-a-large-scale-bias-controlled-dataset-for-pushing-the-limits-of-object-recognition-models>
AUTHORS: Andrei Barbu, David Mayo, Julian Alverio, William Luo, Christopher Wang, Dan Gutfreund, Josh Tenenbaum, Boris Katz
HIGHLIGHT: We collect a large real-world test set, ObjectNet, for object recognition with controls where object backgrounds, rotations, and imaging viewpoints are random.

848, TITLE: Certified Adversarial Robustness with Additive Noise
<https://papers.nips.cc/paper/9143-certified-adversarial-robustness-with-additive-noise>
AUTHORS: Bai Li, Changyou Chen, Wenlin Wang, Lawrence Carin
HIGHLIGHT: To address these limitations, we introduce a framework that is scalable and provides certified bounds on the norm of the input manipulation for constructing adversarial examples.

849, TITLE: Tight Dimensionality Reduction for Sketching Low Degree Polynomial Kernels
<https://papers.nips.cc/paper/9144-tight-dimensionality-reduction-for-sketching-low-degree-polynomial-kernels>
AUTHORS: Michela Meister, Tamas Sarlos, David Woodruff
HIGHLIGHT: We give a new analysis of this sketch, providing nearly optimal bounds.

850, TITLE: Non-Cooperative Inverse Reinforcement Learning
<https://papers.nips.cc/paper/9145-non-cooperative-inverse-reinforcement-learning>
AUTHORS: Xiangyuan Zhang, Kaiqing Zhang, Erik Miebling, Tamer Basar
HIGHLIGHT: To describe such strategic situations, we introduce the non-cooperative inverse reinforcement learning (N-CIRL) formalism.

851, TITLE: DINGO: Distributed Newton-Type Method for Gradient-Norm Optimization
<https://papers.nips.cc/paper/9146-dingo-distributed-newton-type-method-for-gradient-norm-optimization>
AUTHORS: Rixon Crane, Fred Roosta
HIGHLIGHT: For optimization of a large sum of functions in a distributed computing environment, we present a novel communication efficient Newton-type algorithm that enjoys a variety of advantages over similar existing methods.

852, TITLE: Sobolev Independence Criterion
<https://papers.nips.cc/paper/9147-sobolev-independence-criterion>
AUTHORS: Youssef Mroueh, Tom Sercu, Mattia Rigotti, Inkit Padhi, Cicero Nogueira dos Santos
HIGHLIGHT: We propose the Sobolev Independence Criterion (SIC), an interpretable dependency measure between a high dimensional random variable X and a response variable Y . SIC decomposes to the sum of feature importance scores and hence can be used for nonlinear feature selection.

853, TITLE: Maximum Entropy Monte-Carlo Planning
<https://papers.nips.cc/paper/9148-maximum-entropy-monte-carlo-planning>
AUTHORS: Chenjun Xiao, Ruitong Huang, Jincheng Mei, Dale Schuurmans, Martin M?ller
HIGHLIGHT: We develop a new algorithm for online planning in large scale sequential decision problems that improves upon the worst case efficiency of UCT.

854, TITLE: Learning from brains how to regularize machines
<https://papers.nips.cc/paper/9149-learning-from-brains-how-to-regularize-machines>
AUTHORS: Zhe Li, Wieland Brendel, Edgar Walker, Erick Cobos, Taliah Muhammad, Jacob Reimer, Matthias Bethge, Fabian Sinz, Zachary Pitkow, Andreas Tolias

HIGHLIGHT: We propose to regularize CNNs using large-scale neuroscience data to learn more robust neural features in terms of representational similarity.

855, **TITLE:** Using Statistics to Automate Stochastic Optimization
<https://papers.nips.cc/paper/9150-using-statistics-to-automate-stochastic-optimization>

AUTHORS: Hunter Lang, Lin Xiao, Pengchuan Zhang

HIGHLIGHT: Rather than changing the learning rate at each iteration, we propose an approach that automates the most common hand-tuning heuristic: use a constant learning rate until "progress stops," then drop.

856, **TITLE:** Zero-shot Knowledge Transfer via Adversarial Belief Matching
<https://papers.nips.cc/paper/9151-zero-shot-knowledge-transfer-via-adversarial-belief-matching>

AUTHORS: Paul Micaelli, Amos J. Storkey

HIGHLIGHT: We propose a novel method which trains a student to match the predictions of its teacher without using any data or metadata.

857, **TITLE:** Differentiable Convex Optimization Layers
<https://papers.nips.cc/paper/9152-differentiable-convex-optimization-layers>

AUTHORS: Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, J. Zico Kolter

HIGHLIGHT: In this paper, we propose an approach to differentiating through disciplined convex programs, a subclass of convex optimization problems used by domain-specific languages (DSLs) for convex optimization.

858, **TITLE:** Random Tessellation Forests
<https://papers.nips.cc/paper/9153-random-tessellation-forests>

AUTHORS: Shufei Ge, Shijia Wang, Yee Whye Teh, Liangliang Wang, Lloyd Elliott

HIGHLIGHT: Motivated by the need for a multi-dimensional partitioning tree with non-axis aligned cuts, we propose the Random Tessellation Process, a framework that includes the Mondrian process as a special case.

859, **TITLE:** Learning Nearest Neighbor Graphs from Noisy Distance Samples
<https://papers.nips.cc/paper/9154-learning-nearest-neighbor-graphs-from-noisy-distance-samples>

AUTHORS: Blake Mason, Ardhendu Tripathy, Robert Nowak

HIGHLIGHT: In this paper, we propose an active algorithm to find the graph with high probability and analyze its query complexity.

860, **TITLE:** Lookahead Optimizer: k steps forward, 1 step back
<https://papers.nips.cc/paper/9155-lookahead-optimizer-k-steps-forward-1-step-back>

AUTHORS: Michael Zhang, James Lucas, Jimmy Ba, Geoffrey E. Hinton

HIGHLIGHT: In this paper, we propose a new optimization algorithm, Lookahead, that is orthogonal to these previous approaches and iteratively updates two sets of weights.

861, **TITLE:** Learning to Predict 3D Objects with an Interpolation-based Differentiable Renderer
<https://papers.nips.cc/paper/9156-learning-to-predict-3d-objects-with-an-interpolation-based-differentiable-renderer>

AUTHORS: Wenzheng Chen, Huan Ling, Jun Gao, Edward Smith, Jaakko Lehtinen, Alec Jacobson, Sanja Fidler

HIGHLIGHT: In this paper, we present DIB-Render, a novel rendering framework through which gradients can be analytically computed.

862, **TITLE:** Covariate-Powered Empirical Bayes Estimation
<https://papers.nips.cc/paper/9157-covariate-powered-empirical-bayes-estimation>

AUTHORS: Nikolaos Ignatiadis, Stefan Wager

HIGHLIGHT: In this paper, we propose a flexible plug-in empirical Bayes estimator that synthesizes both sources of information and may leverage any black-box predictive model.

863, **TITLE:** Understanding the Role of Momentum in Stochastic Gradient Methods
<https://papers.nips.cc/paper/9158-understanding-the-role-of-momentum-in-stochastic-gradient-methods>

AUTHORS: Igor Gitman, Hunter Lang, Pengchuan Zhang, Lin Xiao

HIGHLIGHT: In this paper, we use the general formulation of QHM to give a unified analysis of several popular algorithms, covering their asymptotic convergence conditions, stability regions, and properties of their stationary distributions.

864, **TITLE:** A neurally plausible model for online recognition and postdiction in a dynamical environment
<https://papers.nips.cc/paper/9159-a-neurally-plausible-model-for-online-recognition-and-postdiction-in-a-dynamical-environment>

AUTHORS: Li Kevin Wenliang, Maneesh Sahani

HIGHLIGHT: Here, we propose a general framework for neural probabilistic inference in dynamic models based on the distributed distributional code (DDC) representation of uncertainty, naturally extending the underlying encoding to incorporate implicit probabilistic beliefs about both present and past.

865, **TITLE:** Guided Meta-Policy Search
<https://papers.nips.cc/paper/9160-guided-meta-policy-search>

AUTHORS: Russell Mendonca, Abhishek Gupta, Rosen Kravev, Pieter Abbeel, Sergey Levine, Chelsea Finn
HIGHLIGHT: To this end, we propose to learn a reinforcement learning procedure in a federated way, where individual off-policy learners can solve the individual meta-training tasks, and then consolidate these solutions into a single meta-learner.

866, **TITLE:** Towards Optimal Off-Policy Evaluation for Reinforcement Learning with Marginalized Importance Sampling
<https://papers.nips.cc/paper/9161-towards-optimal-off-policy-evaluation-for-reinforcement-learning-with-marginalized-importance-sampling>

AUTHORS: Tengyang Xie, Yifei Ma, Yu-Xiang Wang
HIGHLIGHT: Motivated by the many real-world applications of reinforcement learning (RL) that require safe-policy iterations, we consider the problem of off-policy evaluation (OPE) --- the problem of evaluating a new policy using the historical data obtained by different behavior policies --- under the model of nonstationary episodic Markov Decision Processes (MDP) with a long horizon and a large action space.

867, **TITLE:** Contextual Bandits with Cross-Learning
<https://papers.nips.cc/paper/9162-contextual-bandits-with-cross-learning>

AUTHORS: Santiago Balseiro, Negin Golrezaei, Mohammad Mahdian, Vahab Mirrokni, Jon Schneider
HIGHLIGHT: We demonstrate algorithms for the contextual bandits problem with cross-learning that remove the dependence on \mathcal{S} and achieve regret $\tilde{O}(\sqrt{KT})$ (when contexts are stochastic with known distribution), $\tilde{O}(K^{1/3}T^{2/3})$ (when contexts are stochastic with unknown distribution), and $\tilde{O}(\sqrt{KT})$ (when contexts are adversarial but rewards are stochastic).

868, **TITLE:** Evaluating Protein Transfer Learning with TAPE
<https://papers.nips.cc/paper/9163-evaluating-protein-transfer-learning-with-tape>

AUTHORS: Roshan Rao, Nicholas Bhattacharya, Neil Thomas, Yan Duan, Peter Chen, John Canny, Pieter Abbeel, Yun Song
HIGHLIGHT: To facilitate progress in this field, we introduce the Tasks Assessing Protein Embeddings (TAPE), a set of five biologically relevant semi-supervised learning tasks spread across different domains of protein biology.

869, **TITLE:** A Bayesian Theory of Conformity in Collective Decision Making
<https://papers.nips.cc/paper/9164-a-bayesian-theory-of-conformity-in-collective-decision-making>

AUTHORS: Koosha Khalvati, Saghar Mirbagheri, Seongmin A. Park, Jean-Claude Dreher, Rajesh PN Rao
HIGHLIGHT: In this paper, we present a new Bayesian theory of collective decision making based on a simple yet most commonly observed behavior: conformity.

870, **TITLE:** Regularization Matters: Generalization and Optimization of Neural Nets v.s. their Induced Kernel
<https://papers.nips.cc/paper/9165-regularization-matters-generalization-and-optimization-of-neural-nets-vs-their-induced-kernel>

AUTHORS: Colin Wei, Jason D. Lee, Qiang Liu, Tengyu Ma
HIGHLIGHT: We show that sample efficiency can indeed depend on the presence of the regularizer: we construct a simple distribution in d dimensions which the optimal regularized neural net learns with $\tilde{O}(d)$ samples but the NTK requires $\tilde{O}(d^2)$ samples to learn.

871, **TITLE:** Data-dependent Sample Complexity of Deep Neural Networks via Lipschitz Augmentation
<https://papers.nips.cc/paper/9166-data-dependent-sample-complexity-of-deep-neural-networks-via-lipschitz-augmentation>

AUTHORS: Colin Wei, Tengyu Ma
HIGHLIGHT: For feedforward neural nets as well as RNNs, we obtain tighter Rademacher complexity bounds by considering additional data-dependent properties of the network: the norms of the hidden layers of the network, and the norms of the Jacobians of each layer with respect to all previous layers.

872, **TITLE:** A Benchmark for Interpretability Methods in Deep Neural Networks
<https://papers.nips.cc/paper/9167-a-benchmark-for-interpretability-methods-in-deep-neural-networks>

AUTHORS: Sara Hooker, Dumitru Erhan, Pieter-Jan Kindermans, Been Kim
HIGHLIGHT: We propose an empirical measure of the approximate accuracy of feature importance estimates in deep neural networks.

873, **TITLE:** Memory Efficient Adaptive Optimization
<https://papers.nips.cc/paper/9168-memory-efficient-adaptive-optimization>

- AUTHORS: Rohan Anil, Vineet Gupta, Tomer Koren, Yoram Singer
HIGHLIGHT: We describe an effective and flexible adaptive optimization method with greatly reduced memory overhead.
- 874, TITLE: Dynamic Incentive-Aware Learning: Robust Pricing in Contextual Auctions
<https://papers.nips.cc/paper/9169-dynamic-incentive-aware-learning-robust-pricing-in-contextual-auctions>
AUTHORS: Negin Golrezaei, Adel Javanmard, Vahab Mirrokni
HIGHLIGHT: We propose two learning policies that are robust to such strategic behavior.
- 875, TITLE: Convergence-Rate-Matching Discretization of Accelerated Optimization Flows Through Opportunistic State-Triggered Control
<https://papers.nips.cc/paper/9170-convergence-rate-matching-discretization-of-accelerated-optimization-flows-through-opportunistic-state-triggered-control>
AUTHORS: Miguel Vaquero, Jorge Cortes
HIGHLIGHT: This paper provides a novel approach through the idea of opportunistic state-triggered control.
- 876, TITLE: A Unified Framework for Data Poisoning Attack to Graph-based Semi-supervised Learning
<https://papers.nips.cc/paper/9171-a-unified-framework-for-data-poisoning-attack-to-graph-based-semi-supervised-learning>
AUTHORS: Xuanqing Liu, Si Si, Jerry Zhu, Yang Li, Cho-Jui Hsieh
HIGHLIGHT: In this paper, we proposed a general framework for data poisoning attacks to graph-based semi-supervised learning (G-SSL).
- 877, TITLE: Compositional generalization through meta sequence-to-sequence learning
<https://papers.nips.cc/paper/9172-compositional-generalization-through-meta-sequence-to-sequence-learning>
AUTHORS: Brenden M. Lake
HIGHLIGHT: In this paper, I show how memory-augmented neural networks can be trained to generalize compositionally through meta seq2seq learning.
- 878, TITLE: Bayesian Joint Estimation of Multiple Graphical Models
<https://papers.nips.cc/paper/9173-bayesian-joint-estimation-of-multiple-graphical-models>
AUTHORS: Lingrui Gan, Xinming Yang, Naveen Narisetty, Feng Liang
HIGHLIGHT: In this paper, we propose a novel Bayesian group regularization method based on the spike and slab Lasso priors for jointly estimating multiple graphical models.
- 879, TITLE: Practical Two-Step Lookahead Bayesian Optimization
<https://papers.nips.cc/paper/9174-practical-two-step-lookahead-bayesian-optimization>
AUTHORS: Jian Wu, Peter Frazier
HIGHLIGHT: This paper proposes a computationally efficient algorithm that provides an accurate solution to the two-step lookahead Bayesian optimization problem in seconds to at most several minutes of computation per batch of evaluations.
- 880, TITLE: Leader Stochastic Gradient Descent for Distributed Training of Deep Learning Models
<https://papers.nips.cc/paper/9175-leader-stochastic-gradient-descent-for-distributed-training-of-deep-learning-models>
AUTHORS: Yunfei Teng, Wenbo Gao, Fran?ois Chalus, Anna E. Choromanska, Donald Goldfarb, Adrian Weller
HIGHLIGHT: We propose a new algorithm, whose parameter updates rely on two forces: a regular gradient step, and a corrective direction dictated by the currently best-performing worker (leader).
- 881, TITLE: A Convex Relaxation Barrier to Tight Robustness Verification of Neural Networks
<https://papers.nips.cc/paper/9176-a-convex-relaxation-barrier-to-tight-robustness-verification-of-neural-networks>
AUTHORS: Hadi Salman, Greg Yang, Huan Zhang, Cho-Jui Hsieh, Pengchuan Zhang
HIGHLIGHT: In this paper, we unify all existing LP-relaxed verifiers, to the best of our knowledge, under a general convex relaxation framework.
- 882, TITLE: Neural Jump Stochastic Differential Equations
<https://papers.nips.cc/paper/9177-neural-jump-stochastic-differential-equations>
AUTHORS: Junteng Jia, Austin R. Benson
HIGHLIGHT: To this end, we introduce Neural Jump Stochastic Differential Equations that provide a data-driven approach to learn continuous and discrete dynamic behavior, i.e., hybrid systems that both flow and jump.
- 883, TITLE: Learning metrics for persistence-based summaries and applications for graph classification
<https://papers.nips.cc/paper/9178-learning-metrics-for-persistence-based-summaries-and-applications-for-graph-classification>
AUTHORS: Qi Zhao, Yusu Wang

HIGHLIGHT: We study this problem and develop a new weighted kernel, called WKPI, for persistence summaries, as well as an optimization framework to learn the weight (and thus kernel).

884, TITLE: On the Value of Target Data in Transfer Learning
<https://papers.nips.cc/paper/9179-on-the-value-of-target-data-in-transfer-learning>

AUTHORS: Steve Hanneke, Samory Kpotufe

HIGHLIGHT: We aim to understand the value of additional labeled or unlabeled target data in transfer learning, for any given amount of source data; this is motivated by practical questions around minimizing sampling costs, whereby, target data is usually harder or costlier to acquire than source data, but can yield better accuracy.

885, TITLE: Stochastic Variance Reduced Primal Dual Algorithms for Empirical Composition Optimization
<https://papers.nips.cc/paper/9180-stochastic-variance-reduced-primal-dual-algorithms-for-empirical-composition-optimization>

AUTHORS: Adithya M Devraj, Jianshu Chen

HIGHLIGHT: We take a novel approach to solving this problem by reformulating the original minimization objective into an equivalent min-max objective, which brings out all the empirical averages that are originally inside the nonlinear loss functions.

886, TITLE: On Robustness of Principal Component Regression
<https://papers.nips.cc/paper/9181-on-robustness-of-principal-component-regression>

AUTHORS: Anish Agarwal, Devavrat Shah, Dennis Shen, Dogyoon Song

HIGHLIGHT: As the main contribution of this work, we address this challenge by rigorously establishing that PCR is robust to noisy, sparse, and possibly mixed valued covariates.

887, TITLE: Meta Learning with Relational Information for Short Sequences
<https://papers.nips.cc/paper/9182-meta-learning-with-relational-information-for-short-sequences>

AUTHORS: Yujia Xie, Haoming Jiang, Feng Liu, Tuo Zhao, Hongyuan Zha

HIGHLIGHT: This paper proposes a new meta-learning method -- named HARMLESS (HAWkes Relational Meta Learning method for Short Sequences) for learning heterogeneous point process models from a collection of short event sequence data along with a relational network.

888, TITLE: Residual Flows for Invertible Generative Modeling
<https://papers.nips.cc/paper/9183-residual-flows-for-invertible-generative-modeling>

AUTHORS: Tian Qi Chen, Jens Behrmann, David K. Duvenaud, Joern-Henrik Jacobsen

HIGHLIGHT: We give a tractable unbiased estimate of the log density, and reduce the memory required during training by a factor of ten.

889, TITLE: Multi-Agent Common Knowledge Reinforcement Learning
<https://papers.nips.cc/paper/9184-multi-agent-common-knowledge-reinforcement-learning>

AUTHORS: Christian Schroeder de Witt, Jakob Foerster, Gregory Farquhar, Philip Torr, Wendelin Boehmer, Shimon Whiteson

HIGHLIGHT: In this paper, we show that common knowledge between agents allows for complex decentralised coordination.

890, TITLE: Learning to Learn By Self-Critique
<https://papers.nips.cc/paper/9185-learning-to-learn-by-self-critique>

AUTHORS: Antreas Antoniou, Amos J. Storkey

HIGHLIGHT: In this paper, we propose a framework called `\emph{Self-Critique and Adapt}` or SCA.

891, TITLE: Wide Feedforward or Recurrent Neural Networks of Any Architecture are Gaussian Processes
<https://papers.nips.cc/paper/9186-wide-feedforward-or-recurrent-neural-networks-of-any-architecture-are-gaussian-processes>

AUTHORS: Greg Yang

HIGHLIGHT: More generally, we introduce a language for expressing neural network computations, and our result encompasses all such expressible neural networks.

892, TITLE: Neural Networks with Cheap Differential Operators
<https://papers.nips.cc/paper/9187-neural-networks-with-cheap-differential-operators>

AUTHORS: Tian Qi Chen, David K. Duvenaud

HIGHLIGHT: We describe a family of neural network architectures that allow easy access to a family of differential operators involving `\emph{dimension-wise derivatives}`, and we show how to modify the backward computation graph to compute them efficiently.

893, TITLE: Transductive Zero-Shot Learning with Visual Structure Constraint

- <https://papers.nips.cc/paper/9188-transductive-zero-shot-learning-with-visual-structure-constraint>
AUTHORS: Ziyu Wan, Dongdong Chen, Yan Li, Xingguang Yan, Junge Zhang, Yizhou Yu, Jing Liao
HIGHLIGHT: Based on the observation that visual features of test instances can be separated into different clusters, we propose a new visual structure constraint on class centers for transductive ZSL, to improve the generality of the projection function (ie alleviate the above domain shift problem).
- 894, TITLE: Dying Experts: Efficient Algorithms with Optimal Regret Bounds
<https://papers.nips.cc/paper/9189-dying-experts-efficient-algorithms-with-optimal-regret-bounds>
AUTHORS: Hamid Shayestehmanesh, Sajjad Azami, Nishant A. Mehta
HIGHLIGHT: We study a variant of decision-theoretic online learning in which the set of experts that are available to Learner can shrink over time.
- 895, TITLE: Model Similarity Mitigates Test Set Overuse
<https://papers.nips.cc/paper/9190-model-similarity-mitigates-test-set-overuse>
AUTHORS: Horia Mania, John Miller, Ludwig Schmidt, Moritz Hardt, Benjamin Recht
HIGHLIGHT: We proffer a new explanation for the apparent longevity of test data: Many proposed models are similar in their predictions and we prove that this similarity mitigates overfitting.
- 896, TITLE: A unified theory for the origin of grid cells through the lens of pattern formation
<https://papers.nips.cc/paper/9191-a-unified-theory-for-the-origin-of-grid-cells-through-the-lens-of-pattern-formation>
AUTHORS: Ben Sorscher, Gabriel Mel, Surya Ganguli, Samuel Ocko
HIGHLIGHT: Here we provide an analytic theory that unifies the two perspectives by casting the learning dynamics of neural networks trained on navigational tasks as a pattern forming dynamical system.
- 897, TITLE: On Sample Complexity Upper and Lower Bounds for Exact Ranking from Noisy Comparisons
<https://papers.nips.cc/paper/9192-on-sample-complexity-upper-and-lower-bounds-for-exact-ranking-from-noisy-comparisons>
AUTHORS: Wenbo Ren, Jia (Kevin) Liu, Ness Shroff
HIGHLIGHT: Different from most previous works, in this paper, we have three main novelties: (i) compared to prior works, our upper bounds (algorithms) and lower bounds on the sample complexity (aka number of comparisons) require the minimal assumptions on the instances, and are not restricted to specific models; (ii) we give lower bounds and upper bounds on instances with \neq noise levels; and (iii) this paper aims at the exact ranking without knowledge on the instances, while most of the previous works either focus on approximate rankings or study exact ranking but require prior knowledge.
- 898, TITLE: Hierarchical Decision Making by Generating and Following Natural Language Instructions
<https://papers.nips.cc/paper/9193-hierarchical-decision-making-by-generating-and-following-natural-language-instructions>
AUTHORS: Hengyuan Hu, Denis Yarats, Qucheng Gong, Yuandong Tian, Mike Lewis
HIGHLIGHT: We introduce a challenging real-time strategy game environment in which the actions of a large number of units must be coordinated across long time scales.
- 899, TITLE: SHE: A Fast and Accurate Deep Neural Network for Encrypted Data
<https://papers.nips.cc/paper/9194-she-a-fast-and-accurate-deep-neural-network-for-encrypted-data>
AUTHORS: Qian Lou, Lei Jiang
HIGHLIGHT: In this paper, we propose a Shift-accumulation-based LHE-enabled deep neural network (SHE) for fast and accurate inferences on encrypted data.
- 900, TITLE: Locality-Sensitive Hashing for f-Divergences: Mutual Information Loss and Beyond
<https://papers.nips.cc/paper/9195-locality-sensitive-hashing-for-f-divergences-mutual-information-loss-and-beyond>
AUTHORS: Lin Chen, Hossein Esfandiari, Gang Fu, Vahab Mirrokni
HIGHLIGHT: In this paper, we aim to develop LSH schemes for distance functions that measure the distance between two probability distributions, particularly for f-divergences as well as a generalization to capture mutual information loss.
- 901, TITLE: A Game Theoretic Approach to Class-wise Selective Rationalization
<https://papers.nips.cc/paper/9196-a-game-theoretic-approach-to-class-wise-selective-rationalization>
AUTHORS: Shiyu Chang, Yang Zhang, Mo Yu, Tommi Jaakkola
HIGHLIGHT: To this end, we propose a new game theoretic approach to class-dependent rationalization, where the method is specifically trained to highlight evidence supporting alternative conclusions.
- 902, TITLE: Efficiently avoiding saddle points with zero order methods: No gradients required
<https://papers.nips.cc/paper/9197-efficiently-avoiding-saddle-points-with-zero-order-methods-no-gradients-required>
AUTHORS: Emmanouil-Vasilios Vlastakis-Gkaragkounis, Lampros Flokas, Georgios Piliouras

HIGHLIGHT: We consider the case of derivative-free algorithms for non-convex optimization, also known as zero order algorithms, that use only function evaluations rather than gradients

903, TITLE: Metamers of neural networks reveal divergence from human perceptual systems
<https://papers.nips.cc/paper/9198-metamers-of-neural-networks-reveal-divergence-from-human-perceptual-systems>
AUTHORS: Jenelle Feather, Alex Durango, Ray Gonzalez, Josh McDermott
HIGHLIGHT: To more thoroughly investigate their similarity to biological systems, we synthesized model metamers – stimuli that produce the same responses at some stage of a network’s representation.

904, TITLE: Spatial-Aware Feature Aggregation for Image based Cross-View Geo-Localization
<https://papers.nips.cc/paper/9199-spatial-aware-feature-aggregation-for-image-based-cross-view-geo-localization>
AUTHORS: Yujiao Shi, Liu Liu, Xin Yu, Hongdong Li
HIGHLIGHT: In this paper, we develop a new deep network to explicitly address these inherent differences between ground and aerial views.

905, TITLE: Decentralized sketching of low rank matrices
<https://papers.nips.cc/paper/9200-decentralized-sketching-of-low-rank-matrices>
AUTHORS: Rakshith Sharma Srinivasa, Kiryung Lee, Marius Junge, Justin Romberg
HIGHLIGHT: Leveraging the joint structure between the columns, we propose a method to recover the matrix to within an epsilon relative error in the Frobenius norm from a total of $O(r(d_1 + d_2)\log^6(d_1 + d_2)/\epsilon^2)$ observations.

906, TITLE: Average Case Column Subset Selection for Entrywise ℓ_1 -Norm Loss
https://papers.nips.cc/paper/9201-average-case-column-subset-selection-for-entrywise-ell_1-norm-loss
AUTHORS: Zhao Song, David Woodruff, Peilin Zhong
HIGHLIGHT: We study the column subset selection problem with respect to the entrywise ℓ_1 -norm loss.

907, TITLE: Efficient Forward Architecture Search
<https://papers.nips.cc/paper/9202-efficient-forward-architecture-search>
AUTHORS: Hanzhang Hu, John Langford, Rich Caruana, Saurajit Mukherjee, Eric J. Horvitz, Debadeepta Dey
HIGHLIGHT: We propose a neural architecture search (NAS) algorithm, Petridish, to iteratively add shortcut connections to existing network layers.

908, TITLE: Unsupervised Meta-Learning for Few-Shot Image Classification
<https://papers.nips.cc/paper/9203-unsupervised-meta-learning-for-few-shot-image-classification>
AUTHORS: Siavash Khodadadeh, Ladislau Boloni, Mubarak Shah
HIGHLIGHT: In this paper, we propose UMTRA, an algorithm that performs unsupervised, model-agnostic meta-learning for classification tasks.

909, TITLE: Learning Mixtures of Plackett-Luce Models from Structured Partial Orders
<https://papers.nips.cc/paper/9204-learning-mixtures-of-plackett-luce-models-from-structured-partial-orders>
AUTHORS: Zhibing Zhao, Lirong Xia
HIGHLIGHT: In this paper, we focus on three popular structures of partial orders: ranked top- ℓ_1 , ℓ_2 -way, and choice data over a subset of alternatives.

910, TITLE: Certainty Equivalence is Efficient for Linear Quadratic Control
<https://papers.nips.cc/paper/9205-certainty-equivalence-is-efficient-for-linear-quadratic-control>
AUTHORS: Horia Mania, Stephen Tu, Benjamin Recht
HIGHLIGHT: We show that for both the fully and partially observed settings, the sub-optimality gap between the cost incurred by playing the certainty equivalent controller on the true system and the cost incurred by using the optimal LQ controller enjoys a fast statistical rate, scaling as the square of the parameter error.

911, TITLE: Scalable Bayesian inference of dendritic voltage via spatiotemporal recurrent state space models
<https://papers.nips.cc/paper/9206-scalable-bayesian-inference-of-dendritic-voltage-via-spatiotemporal-recurrent-state-space-models>
AUTHORS: Ruoxi Sun, Ian Kinsella, Scott Linderman, Liam Paninski
HIGHLIGHT: Here we introduce a scalable fully Bayesian approach.

912, TITLE: Logarithmic Regret for Online Control
<https://papers.nips.cc/paper/9207-logarithmic-regret-for-online-control>
AUTHORS: Naman Agarwal, Elad Hazan, Karan Singh

HIGHLIGHT: We study optimal regret bounds for control in linear dynamical systems under adversarially changing strongly convex cost functions, given the knowledge of transition dynamics.

913, **TITLE:** Elliptical Perturbations for Differential Privacy
<https://papers.nips.cc/paper/9208-elliptical-perturbations-for-differential-privacy>

AUTHORS: Matthew Reimherr, Jordan Awan

HIGHLIGHT: We study elliptical distributions in locally convex vector spaces, and determine conditions when they can or cannot be used to satisfy differential privacy (DP).

914, **TITLE:** Devign: Effective Vulnerability Identification by Learning Comprehensive Program Semantics via Graph Neural Networks
<https://papers.nips.cc/paper/9209-devign-effective-vulnerability-identification-by-learning-comprehensive-program-semantics-via-graph-neural-networks>

AUTHORS: Yaqin Zhou, Shangqing Liu, Jingkai Siow, Xiaoning Du, Yang Liu

HIGHLIGHT: Inspired by the work on manually-defined patterns of vulnerabilities from various code representation graphs and the recent advance on graph neural networks, we propose Devign, a general graph neural network based model for graph-level classification through learning on a rich set of code semantic representations.

915, **TITLE:** KNG: The K-Norm Gradient Mechanism
<https://papers.nips.cc/paper/9210-kng-the-k-norm-gradient-mechanism>

AUTHORS: Matthew Reimherr, Jordan Awan

HIGHLIGHT: This paper presents a new mechanism for producing sanitized statistical summaries that achieve ϵ -differential privacy, called the K -Norm Gradient Mechanism, or KNG.

916, **TITLE:** CXPlain: Causal Explanations for Model Interpretation under Uncertainty
<https://papers.nips.cc/paper/9211-explain-causal-explanations-for-model-interpretation-under-uncertainty>

AUTHORS: Patrick Schwab, Walter Karlen

HIGHLIGHT: We present experiments that demonstrate that CXPlain is significantly more accurate and faster than existing model-agnostic methods for estimating feature importance.

917, **TITLE:** Regularized Anderson Acceleration for Off-Policy Deep Reinforcement Learning
<https://papers.nips.cc/paper/9212-regularized-anderson-acceleration-for-off-policy-deep-reinforcement-learning>

AUTHORS: Wenjie Shi, Shiji Song, Hui Wu, Ya-Chu Hsu, Cheng Wu, Gao Huang

HIGHLIGHT: To tackle this problem, we propose a general acceleration method for model-free, off-policy deep RL algorithms by drawing the idea underlying regularized Anderson acceleration (RAA), which is an effective approach to accelerating the solving of fixed point problems with perturbations.

918, **TITLE:** STREETS: A Novel Camera Network Dataset for Traffic Flow
<https://papers.nips.cc/paper/9213-streets-a-novel-camera-network-dataset-for-traffic-flow>

AUTHORS: Corey Snyder, Minh Do

HIGHLIGHT: In this paper, we introduce STREETS, a novel traffic flow dataset from publicly available web cameras in the suburbs of Chicago, IL.

919, **TITLE:** Sequential Neural Processes
<https://papers.nips.cc/paper/9214-sequential-neural-processes>

AUTHORS: Gautam Singh, Jaesik Yoon, Youngsung Son, Sungjin Ahn

HIGHLIGHT: In this paper, we propose Sequential Neural Processes (SNP) which incorporates a temporal state-transition model of stochastic processes and thus extends its modeling capabilities to dynamic stochastic processes.

920, **TITLE:** Policy Continuation with Hindsight Inverse Dynamics
<https://papers.nips.cc/paper/9215-policy-continuation-with-hindsight-inverse-dynamics>

AUTHORS: Hao Sun, Zhizhong Li, Xiaotong Liu, Bolei Zhou, Dahua Lin

HIGHLIGHT: To tackle this difficulty, we propose a new approach called Policy Continuation with Hindsight Inverse Dynamics (PCHID).

921, **TITLE:** Learning to Self-Train for Semi-Supervised Few-Shot Classification
<https://papers.nips.cc/paper/9216-learning-to-self-train-for-semi-supervised-few-shot-classification>

AUTHORS: Xinzhe Li, Qianru Sun, Yaoyao Liu, Qin Zhou, Shibao Zheng, Tat-Seng Chua, Bernt Schiele

HIGHLIGHT: In this paper we propose a novel semi-supervised meta-learning method called learning to self-train (LST) that leverages unlabeled data and specifically meta-learns how to cherry-pick and label such unsupervised data to further improve performance.

- 922, TITLE: Temporal FiLM: Capturing Long-Range Sequence Dependencies with Feature-Wise Modulations.
<https://papers.nips.cc/paper/9217-temporal-film-capturing-long-range-sequence-dependencies-with-feature-wise-modulations>
AUTHORS: Sawyer Bimbaum, Volodymyr Kuleshov, Zayd Enam, Pang Wei W. Koh, Stefano Ermon
HIGHLIGHT: Here, we propose Temporal Feature-Wise Linear Modulation (TFiLM) --- a novel architectural component inspired by adaptive batch normalization and its extensions --- that uses a recurrent neural network to alter the activations of a convolutional model.
- 923, TITLE: From Complexity to Simplicity: Adaptive ES-Active Subspaces for Blackbox Optimization
<https://papers.nips.cc/paper/9218-from-complexity-to-simplicity-adaptive-es-active-subspaces-for-blackbox-optimization>
AUTHORS: Krzysztof M. Choromanski, Aldo Pacchiano, Jack Parker-Holder, Yunhao Tang, Vikas Sindhwani
HIGHLIGHT: We present a new algorithm (ASEBO) for optimizing high-dimensional blackbox functions.
- 924, TITLE: On the Expressive Power of Deep Polynomial Neural Networks
<https://papers.nips.cc/paper/9219-on-the-expressive-power-of-deep-polynomial-neural-networks>
AUTHORS: Joe Kileel, Matthew Trager, Joan Bruna
HIGHLIGHT: This paper proposes the dimension of this variety as a precise measure of the expressive power of polynomial neural networks.
- 925, TITLE: DETOX: A Redundancy-based Framework for Faster and More Robust Gradient Aggregation
<https://papers.nips.cc/paper/9220-detox-a-redundancy-based-framework-for-faster-and-more-robust-gradient-aggregation>
AUTHORS: Shashank Rajput, Hongyi Wang, Zachary Charles, Dimitris Papailiopoulos
HIGHLIGHT: In this work, we present DETOX, a Byzantine-resilient distributed training framework that combines algorithmic redundancy with robust aggregation.
- 926, TITLE: Can SGD Learn Recurrent Neural Networks with Provable Generalization?
<https://papers.nips.cc/paper/9221-can-sgd-learn-recurrent-neural-networks-with-provable-generalization>
AUTHORS: Zeyuan Allen-Zhu, Yuanzhi Li
HIGHLIGHT: In this paper, we show using the vanilla stochastic gradient descent (SGD), RNN can actually learn some notable concept class $\text{emph}\{\text{efficiently}\}$, meaning that both time and sample complexity scale $\text{emph}\{\text{polynomially}\}$ in the input length (or almost polynomially, depending on the concept).
- 927, TITLE: Limits of Private Learning with Access to Public Data
<https://papers.nips.cc/paper/9222-limits-of-private-learning-with-access-to-public-data>
AUTHORS: Raef Bassily, Shay Moran, Noga Alon
HIGHLIGHT: We consider learning problems where the training set consists of two types of examples: private and public. The goal is to design a learning algorithm that satisfies differential privacy only with respect to the private examples.
- 928, TITLE: Discrete Object Generation with Reversible Inductive Construction
<https://papers.nips.cc/paper/9223-discrete-object-generation-with-reversible-inductive-construction>
AUTHORS: Ari Seff, Wenda Zhou, Farhan Damani, Abigail Doyle, Ryan P. Adams
HIGHLIGHT: Here, we present a generative model for discrete objects employing a Markov chain where transitions are restricted to a set of local operations that preserve validity.
- 929, TITLE: Efficient Near-Optimal Testing of Community Changes in Balanced Stochastic Block Models
<https://papers.nips.cc/paper/9224-efficient-near-optimal-testing-of-community-changes-in-balanced-stochastic-block-models>
AUTHORS: Aditya Gangrade, Praveen Venkatesh, Bobak Nazer, Venkatesh Saligrama
HIGHLIGHT: We propose and analyze the problems of $\text{textit}\{\text{community goodness-of-fit and two-sample testing}\}$ for stochastic block models (SBM), where changes arise due to modification in community memberships of nodes.
- 930, TITLE: Keeping Your Distance: Solving Sparse Reward Tasks Using Self-Balancing Shaped Rewards
<https://papers.nips.cc/paper/9225-keeping-your-distance-solving-sparse-reward-tasks-using-self-balancing-shaped-rewards>
AUTHORS: Alexander Trott, Stephan Zheng, Caiming Xiong, Richard Socher
HIGHLIGHT: We introduce a simple and effective model-free method to learn from shaped distance-to-goal rewards on tasks where success depends on reaching a goal state.
- 931, TITLE: Superset Technique for Approximate Recovery in One-Bit Compressed Sensing
<https://papers.nips.cc/paper/9226-superset-technique-for-approximate-recovery-in-one-bit-compressed-sensing>
AUTHORS: Larkin Flodin, Venkata Gandikota, Arya Mazumdar

HIGHLIGHT: In this paper, we propose a generic approach for signal recovery in nonadaptive 1bCS that leads to improved sample complexity for approximate recovery for a variety of signal models, including nonnegative signals and binary signals.

932, **TITLE:** Bandits with Feedback Graphs and Switching Costs

<https://papers.nips.cc/paper/9227-bandits-with-feedback-graphs-and-switching-costs>

AUTHORS: Raman Arora, Teodor Vanislavov Marinov, Mehryar Mohri

HIGHLIGHT: We study the adversarial multi-armed bandit problem where the learner is supplied with partial observations modeled by a feedback graph and where shifting to a new action incurs a fixed switching cost.

933, **TITLE:** Functional Adversarial Attacks

<https://papers.nips.cc/paper/9228-functional-adversarial-attacks>

AUTHORS: Cassidy Laidlaw, Soheil Feizi

HIGHLIGHT: We propose functional adversarial attacks, a novel class of threat models for crafting adversarial examples to fool machine learning models.

934, **TITLE:** Statistical-Computational Tradeoff in Single Index Models

<https://papers.nips.cc/paper/9229-statistical-computational-tradeoff-in-single-index-models>

AUTHORS: Lingxiao Wang, Zhuoran Yang, Zhaoran Wang

HIGHLIGHT: In this paper, we investigate the case when this critical assumption fails to hold, where the problem becomes considerably harder.

935, **TITLE:** On Fenchel Mini-Max Learning

<https://papers.nips.cc/paper/9230-on-fenchel-mini-max-learning>

AUTHORS: Chenyang Tao, Liqun Chen, Shuyang Dai, Junya Chen, Ke Bai, Dong Wang, Jianfeng Feng, Wenlian Lu, Georgiy Bobashev, Lawrence Carin

HIGHLIGHT: We present a novel probabilistic learning framework, called Fenchel Mini-Max Learning (FML), that accommodates all four desiderata in a flexible and scalable manner.

936, **TITLE:** MarginGAN: Adversarial Training in Semi-Supervised Learning

<https://papers.nips.cc/paper/9231-margingan-adversarial-training-in-semi-supervised-learning>

AUTHORS: Jinhao Dong, Tong Lin

HIGHLIGHT: Our method is motivated by the success of large-margin classifiers and the recent viewpoint that good semi-supervised learning requires a bad GAN.

937, **TITLE:** Poincare Recurrence, Cycles and Spurious Equilibria in Gradient-Descent-Ascent for Non-Convex Non-Concave Zero-Sum Games

<https://papers.nips.cc/paper/9232-poincare-recurrence-cycles-and-spurious-equilibria-in-gradient-descent-ascent-for-non-convex-non-concave-zero-sum-games>

AUTHORS: Emmanouil Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, Georgios Piliouras

HIGHLIGHT: We study a wide class of non-convex non-concave min-max games that generalizes over standard bilinear zero-sum games.

938, **TITLE:** A unified variance-reduced accelerated gradient method for convex optimization

<https://papers.nips.cc/paper/9233-a-unified-variance-reduced-accelerated-gradient-method-for-convex-optimization>

AUTHORS: Guanghui Lan, Zhize Li, Yi Zhou

HIGHLIGHT: We propose a novel randomized incremental gradient algorithm, namely, Variance-Reduced Accelerated Gradient (Varag), for finite-sum optimization.

939, **TITLE:** Nearly Tight Bounds for Robust Proper Learning of Halfspaces with a Margin

<https://papers.nips.cc/paper/9234-nearly-tight-bounds-for-robust-proper-learning-of-halfspaces-with-a-margin>

AUTHORS: Ilias Diakonikolas, Daniel Kane, Pasin Manurangsi

HIGHLIGHT: We study the problem of properly learning large margin halfspaces in the agnostic PAC model.

940, **TITLE:** Same-Cluster Querying for Overlapping Clusters

<https://papers.nips.cc/paper/9235-same-cluster-querying-for-overlapping-clusters>

AUTHORS: Wasim Huleihel, Arya Mazumdar, Muriel Medard, Soumyabrata Pal

HIGHLIGHT: In this paper, we look at the more practical scenario of overlapping clusters, and provide upper bounds (with algorithms) on the sufficient number of queries.

941, **TITLE:** Efficient Convex Relaxations for Streaming PCA

<https://papers.nips.cc/paper/9236-efficient-convex-relaxations-for-streaming-pca>

AUTHORS: Raman Arora, Teodor Vanislavov Marinov

HIGHLIGHT: In this work, we give improved bounds on per iteration cost of mini-batched variants of both MSG and Sell₂-RMSG and arrive at an algorithm with total computational complexity matching that of Oja's algorithm.

942, TITLE: Learning Robust Global Representations by Penalizing Local Predictive Power

<https://papers.nips.cc/paper/9237-learning-robust-global-representations-by-penalizing-local-predictive-power>

AUTHORS: Haoan Wang, Songwei Ge, Zachary Lipton, Eric P. Xing

HIGHLIGHT: This paper proposes a method for training robust convolutional networks by penalizing the predictive power of the local representations learned by earlier layers.

943, TITLE: Unsupervised Curricula for Visual Meta-Reinforcement Learning

<https://papers.nips.cc/paper/9238-unsupervised-curricula-for-visual-meta-reinforcement-learning>

AUTHORS: Allan Jabri, Kyle Hsu, Abhishek Gupta, Ben Eysenbach, Sergey Levine, Chelsea Finn

HIGHLIGHT: We formulate unsupervised meta-RL as information maximization between a latent task variable and the meta-learner's data distribution, and describe a practical instantiation which alternates between integration of recent experience into the task distribution and meta-learning of the updated tasks.

944, TITLE: Sample Complexity of Learning Mixture of Sparse Linear Regressions

<https://papers.nips.cc/paper/9239-sample-complexity-of-learning-mixture-of-sparse-linear-regressions>

AUTHORS: Akshay Krishnamurthy, Arya Mazumdar, Andrew McGregor, Soumyabrata Pal

HIGHLIGHT: In this paper, we consider the case where the signal vectors are sparse; this generalizes the popular compressed sensing paradigm.

945, TITLE: Large Scale Adversarial Representation Learning

<https://papers.nips.cc/paper/9240-large-scale-adversarial-representation-learning>

AUTHORS: Jeff Donahue, Karen Simonyan

HIGHLIGHT: In this work we show that progress in image generation quality translates to substantially improved representation learning performance.

946, TITLE: G2SAT: Learning to Generate SAT Formulas

<https://papers.nips.cc/paper/9241-g2sat-learning-to-generate-sat-formulas>

AUTHORS: Jiaxuan You, Haoze Wu, Clark Barrett, Raghuram Ramanujan, Jure Leskovec

HIGHLIGHT: In this work, we present G2SAT, the first deep generative framework that learns to generate SAT formulas from a given set of input formulas.

947, TITLE: Neural Trust Region/Proximal Policy Optimization Attains Globally Optimal Policy

<https://papers.nips.cc/paper/9242-neural-trust-regionproximal-policy-optimization-attains-globally-optimal-policy>

AUTHORS: Boyi Liu, Qi Cai, Zhuoran Yang, Zhaoran Wang

HIGHLIGHT: In this paper, we prove that a variant of PPO and TRPO equipped with overparametrized neural networks converges to the globally optimal policy at a sublinear rate.

948, TITLE: Dimensionality reduction: theoretical perspective on practical measures

<https://papers.nips.cc/paper/9243-dimensionality-reduction-theoretical-perspective-on-practical-measures>

AUTHORS: Yair Bartal, Nova Fandina, Ofer Neiman

HIGHLIGHT: The goal of this paper is to bridge the gap between theory and practice view-points of metric dimensionality reduction, laying the foundation for a theoretical study of more practically oriented analysis.

949, TITLE: Oracle-Efficient Algorithms for Online Linear Optimization with Bandit Feedback

<https://papers.nips.cc/paper/9244-oracle-efficient-algorithms-for-online-linear-optimization-with-bandit-feedback>

AUTHORS: Shinji Ito, Daisuke Hatano, Hanna Sumita, Kei Takemura, Takuro Fukunaga, Naonori Kakimura, Ken-Ichi Kawarabayashi

HIGHLIGHT: We propose computationally efficient algorithms for $\text{online linear optimization with bandit feedback}$, in which a player chooses an action vector from a given (possibly infinite) set $\mathcal{A} \subseteq \mathbb{R}^d$, and then suffers a loss that can be expressed as a linear function in action vectors.

950, TITLE: Multilabel reductions: what is my loss optimising?

<https://papers.nips.cc/paper/9245-multilabel-reductions-what-is-my-loss-optimising>

AUTHORS:

HIGHLIGHT: In this paper, we study five commonly used reductions, including the one-versus-all reduction, a reduction to multiclass classification, and normalised versions of the same, wherein the contribution of each instance is normalised by the number of relevant labels.

951, **TITLE:** Tight Sample Complexity of Learning One-hidden-layer Convolutional Neural Networks
<https://papers.nips.cc/paper/9246-tight-sample-complexity-of-learning-one-hidden-layer-convolutional-neural-networks>
AUTHORS: Yuan Cao, Quanquan Gu
HIGHLIGHT: We propose a novel algorithm called approximate gradient descent for training CNNs, and show that, with high probability, the proposed algorithm with random initialization grants a linear convergence to the ground-truth parameters up to statistical precision.

952, **TITLE:** Deep Gamblers: Learning to Abstain with Portfolio Theory
<https://papers.nips.cc/paper/9247-deep-gamblers-learning-to-abstain-with-portfolio-theory>
AUTHORS: Ziyin Liu, Zhikang Wang, Paul Pu Liang, Russ R. Salakhutdinov, Louis-Philippe Morency, Masahito Ueda
HIGHLIGHT: Inspired by portfolio theory, we propose a loss function for the selective classification problem based on the doubling rate of gambling.

953, **TITLE:** Two Time-scale Off-Policy TD Learning: Non-asymptotic Analysis over Markovian Samples
<https://papers.nips.cc/paper/9248-two-time-scale-off-policy-td-learning-non-asymptotic-analysis-over-markovian-samples>
AUTHORS: Tengyu Xu, Shaofeng Zou, Yingbin Liang
HIGHLIGHT: In contrast to previous studies that characterized the non-asymptotic convergence rate of TDC only under identical and independently distributed (i.i.d.) data samples, we provide the first non-asymptotic convergence analysis for two time-scale TDC under a non-i.i.d. Markovian sample path and linear function approximation.

954, **TITLE:** Transfer Learning via Minimizing the Performance Gap Between Domains
<https://papers.nips.cc/paper/9249-transfer-learning-via-minimizing-the-performance-gap-between-domains>
AUTHORS: Boyu Wang, Jorge Mendez, Mingbo Cai, Eric Eaton
HIGHLIGHT: To formalize this intuition, we define the performance gap as a measure of the discrepancy between the source and target domains.

955, **TITLE:** Splitting Steepest Descent for Growing Neural Architectures
<https://papers.nips.cc/paper/9250-splitting-steepest-descent-for-growing-neural-architectures>
AUTHORS: Lemeng Wu, Dilin Wang, Qiang Liu
HIGHLIGHT: We develop a progressive training approach for neural networks which adaptively grows the network structure by splitting existing neurons to multiple off-springs.

956, **TITLE:** Sequential Experimental Design for Transductive Linear Bandits
<https://papers.nips.cc/paper/9251-sequential-experimental-design-for-transductive-linear-bandits>
AUTHORS: Lalit Jain, Kevin G. Jamieson, Tanner Fiez, Lillian Ratliff
HIGHLIGHT: In this paper we introduce the pure exploration *transductive linear bandit problem*: given a set of measurement vectors $\mathcal{X} \subset \mathbb{R}^d$, a set of items $\mathcal{Z} \subset \mathbb{R}^d$, a fixed confidence δ , and an unknown vector $\theta^* \in \mathbb{R}^d$, the goal is to infer $\arg\max_{z \in \mathcal{Z}} z^\top \theta^*$ with probability $1 - \delta$ by making as few sequentially chosen noisy measurements of the form $x^\top \theta^*$ as possible.

957, **TITLE:** Time Matters in Regularizing Deep Networks: Weight Decay and Data Augmentation Affect Early Learning Dynamics, Matter Little Near Convergence
<https://papers.nips.cc/paper/9252-time-matters-in-regularizing-deep-networks-weight-decay-and-data-augmentation-affect-early-learning-dynamics-matter-little-near-convergence>
AUTHORS: Aditya Sharad Golatkar, Alessandro Achille, Stefano Soatto
HIGHLIGHT: Deep neural networks (DNNs), however, challenge this view: We show that removing regularization after an initial transient period has little effect on generalization, even if the final loss landscape is the same as if there had been no regularization.

958, **TITLE:** Outlier-Robust High-Dimensional Sparse Estimation via Iterative Filtering
<https://papers.nips.cc/paper/9253-outlier-robust-high-dimensional-sparse-estimation-via-iterative-filtering>
AUTHORS: Ilias Diakonikolas, Daniel Kane, Sushrut Karmalkar, Eric Price, Alistair Stewart
HIGHLIGHT: Specifically, we focus on the fundamental problems of robust sparse mean estimation and robust sparse PCA. We give the first practically viable robust estimators for these problems.

959, **TITLE:** Variational Graph Recurrent Neural Networks

<https://papers.nips.cc/paper/9254-variational-graph-recurrent-neural-networks>
AUTHORS: Ehsan Hajiramezani, Arman Hasanzadeh, Krishna Narayanan, Nick Duffield, Mingyuan Zhou, Xiaoning Qian
HIGHLIGHT: In this paper, we develop a novel hierarchical variational model that introduces additional latent random variables to jointly model the hidden states of a graph recurrent neural network (GRNN) to capture both topology and node attribute changes in dynamic graphs.

960, TITLE: Semi-Implicit Graph Variational Auto-Encoders
<https://papers.nips.cc/paper/9255-semi-implicit-graph-variational-auto-encoders>
AUTHORS: Arman Hasanzadeh, Ehsan Hajiramezani, Krishna Narayanan, Nick Duffield, Mingyuan Zhou, Xiaoning Qian
HIGHLIGHT: Semi-implicit graph variational auto-encoder (SIG-VAE) is proposed to expand the flexibility of variational graph auto-encoders (VGAE) to model graph data.

961, TITLE: Unsupervised Learning of Object Keypoints for Perception and Control
<https://papers.nips.cc/paper/9256-unsupervised-learning-of-object-keypoints-for-perception-and-control>
AUTHORS: Tejas D. Kulkarni, Ankush Gupta, Catalin Ionescu, Sebastian Borgeaud, Malcolm Reynolds, Andrew Zisserman, Volodymyr Mnih
HIGHLIGHT: In this work we aim to learn object representations that are useful for control and reinforcement learning (RL).

962, TITLE: A Model-Based Reinforcement Learning with Adversarial Training for Online Recommendation
<https://papers.nips.cc/paper/9257-a-model-based-reinforcement-learning-with-adversarial-training-for-online-recommendation>
AUTHORS: Xueying Bai, Jian Guan, Hongning Wang
HIGHLIGHT: In this work, we propose a model-based reinforcement learning solution which models the user-agent interaction for offline policy learning via a generative adversarial network.

963, TITLE: Optimizing Generalized Rate Metrics with Three Players
<https://papers.nips.cc/paper/9258-optimizing-generalized-rate-metrics-with-three-players>
AUTHORS: Harikrishna Narasimhan, Andrew Cotter, Maya Gupta
HIGHLIGHT: We present a general framework for solving a large class of learning problems with non-linear functions of classification rates.

964, TITLE: Consistency-based Semi-supervised Learning for Object detection
<https://papers.nips.cc/paper/9259-consistency-based-semi-supervised-learning-for-object-detection>
AUTHORS: Jisoo Jeong, Seungeui Lee, Jeesoo Kim, Nojun Kwak
HIGHLIGHT: To alleviate this problem, we propose a Consistency-based Semi-supervised learning method for object Detection (CSD), which is a way of using consistency constraints as a tool for enhancing detection performance by making full use of available unlabeled data.

965, TITLE: Rates of Convergence for Large-scale Nearest Neighbor Classification
<https://papers.nips.cc/paper/9260-rates-of-convergence-for-large-scale-nearest-neighbor-classification>
AUTHORS: Xingye Qiao, Jiexin Duan, Guang Cheng
HIGHLIGHT: For a large data set which cannot be loaded into the memory of a single machine due to computation, communication, privacy, or ownership limitations, we consider the divide and conquer scheme: the entire data set is divided into small subsamples, on which nearest neighbor predictions are made, and then a final decision is reached by aggregating the predictions on subsamples by majority voting.

966, TITLE: An Embedding Framework for Consistent Polyhedral Surrogates
<https://papers.nips.cc/paper/9261-an-embedding-framework-for-consistent-polyhedral-surrogates>
AUTHORS: Jessica Finocchiaro, Rafael Frongillo, Bo Waggoner
HIGHLIGHT: We formalize and study the natural approach of designing convex surrogate loss functions via embeddings for problems such as classification or ranking.

967, TITLE: Cross-Modal Learning with Adversarial Samples
<https://papers.nips.cc/paper/9262-cross-modal-learning-with-adversarial-samples>
AUTHORS: CHAO LI, Shangqian Gao, Cheng Deng, De Xie, Wei Liu
HIGHLIGHT: In this paper, we propose a novel Cross-Modal correlation Learning with Adversarial samples, namely CMLA, which for the first time presents the existence of adversarial samples in cross-modal data.

968, TITLE: Fast-rate PAC-Bayes Generalization Bounds via Shifted Rademacher Processes
<https://papers.nips.cc/paper/9263-fast-rate-pac-bayes-generalization-bounds-via-shifted-rademacher-processes>
AUTHORS: Jun Yang, Shengyang Sun, Daniel M. Roy

- HIGHLIGHT:** The goal of this paper is to extend this bridge between Rademacher complexity and state-of-the-art PAC-Bayesian theory.
- 969, **TITLE:** Input-Cell Attention Reduces Vanishing Saliency of Recurrent Neural Networks
<https://papers.nips.cc/paper/9264-input-cell-attention-reduces-vanishing-saliency-of-recurrent-neural-networks>
AUTHORS: Aya Abdelsalam Ismail, Mohamed Gunady, Luiz Pessoa, Hector Corrada Bravo, Soheil Feizi
HIGHLIGHT: In this work we analyze saliency-based methods for RNNs, both classical and gated cell architectures.
- 970, **TITLE:** Program Synthesis and Semantic Parsing with Learned Code Idioms
<https://papers.nips.cc/paper/9265-program-synthesis-and-semantic-parsing-with-learned-code-idioms>
AUTHORS: Eui Chul Shin, Miltiadis Allamanis, Marc Brockschmidt, Alex Polozov
HIGHLIGHT: In this work, we present Patois, a system that allows a neural program synthesizer to explicitly interleave high-level and low-level reasoning at every generation step.
- 971, **TITLE:** Generalization Bounds of Stochastic Gradient Descent for Wide and Deep Neural Networks
<https://papers.nips.cc/paper/9266-generalization-bounds-of-stochastic-gradient-descent-for-wide-and-deep-neural-networks>
AUTHORS: Yuan Cao, Quanquan Gu
HIGHLIGHT: We study the training and generalization of deep neural networks (DNNs) in the over-parameterized regime, where the network width (i.e., number of hidden nodes per layer) is much larger than the number of training data points.
- 972, **TITLE:** High-Dimensional Optimization in Adaptive Random Subspaces
<https://papers.nips.cc/paper/9267-high-dimensional-optimization-in-adaptive-random-subspaces>
AUTHORS: Jonathan Lacotte, Mert Pilanci, Marco Pavone
HIGHLIGHT: We propose a new randomized optimization method for high-dimensional problems which can be seen as a generalization of coordinate descent to random subspaces.
- 973, **TITLE:** Random Projections with Asymmetric Quantization
<https://papers.nips.cc/paper/9268-random-projections-with-asymmetric-quantization>
AUTHORS: Xiaoyun Li, Ping Li
HIGHLIGHT: In this paper, we investigate the cosine similarity estimators derived in such setting under the Lloyd-Max (LM) quantization scheme.
- 974, **TITLE:** Superposition of many models into one
<https://papers.nips.cc/paper/9269-superposition-of-many-models-into-one>
AUTHORS: Brian Cheung, Alexander Terekhov, Yubei Chen, Pulkit Agrawal, Bruno Olshausen
HIGHLIGHT: We present a method for storing multiple models within a single set of parameters.
- 975, **TITLE:** Private Testing of Distributions via Sample Permutations
<https://papers.nips.cc/paper/9270-private-testing-of-distributions-via-sample-permutations>
AUTHORS: Maryam Aliakbarpour, Ilias Diakonikolas, Daniel Kane, Ronitt Rubinfeld
HIGHLIGHT: In this paper, we use the framework of property testing to design algorithms to test the properties of the distribution that the data is drawn from with respect to differential privacy.
- 976, **TITLE:** McDiarmid-Type Inequalities for Graph-Dependent Variables and Stability Bounds
<https://papers.nips.cc/paper/9271-mcdiarmid-type-inequalities-for-graph-dependent-variables-and-stability-bounds>
AUTHORS: Rui (Ray) Zhang, Xingwu Liu, Yuyi Wang, Liwei Wang
HIGHLIGHT: We consider learning problems in which examples are dependent and their dependency relation is characterized by a graph.
- 977, **TITLE:** How to Initialize your Network? Robust Initialization for WeightNorm & ResNets
<https://papers.nips.cc/paper/9272-how-to-initialize-your-network-robust-initialization-for-weightnorm-resnets>
AUTHORS: Devansh Arpit, Victor Campos, Yoshua Bengio
HIGHLIGHT: To address these issues, we propose a novel parameter initialization strategy that avoids explosion/vanishment of information across layers for weight normalized networks with and without residual connections.
- 978, **TITLE:** On Making Stochastic Classifiers Deterministic
<https://papers.nips.cc/paper/9273-on-making-stochastic-classifiers-deterministic>
AUTHORS: Andrew Cotter, Maya Gupta, Harikrishna Narasimhan
HIGHLIGHT: In this paper, we attempt to answer the theoretical question of how well a stochastic classifier can be approximated by a deterministic one, and compare several different approaches, proving lower and upper bounds.

- 979, TITLE: Statistical Analysis of Nearest Neighbor Methods for Anomaly Detection
<https://papers.nips.cc/paper/9274-statistical-analysis-of-nearest-neighbor-methods-for-anomaly-detection>
AUTHORS: Xiaoyi Gu, Leman Akoglu, Alessandro Rinaldo
HIGHLIGHT: In this paper we are concerned with investigating the performance of NN-based methods for anomaly detection.
- 980, TITLE: Improving Black-box Adversarial Attacks with a Transfer-based Prior
<https://papers.nips.cc/paper/9275-improving-black-box-adversarial-attacks-with-a-transfer-based-prior>
AUTHORS: Shuyu Cheng, Yinpeng Dong, Tianyu Pang, Hang Su, Jun Zhu
HIGHLIGHT: To address these problems, we propose a prior-guided random gradient-free (P-RGF) method to improve black-box adversarial attacks, which takes the advantage of a transfer-based prior and the query information simultaneously.
- 981, TITLE: Break the Ceiling: Stronger Multi-scale Deep Graph Convolutional Networks
<https://papers.nips.cc/paper/9276-break-the-ceiling-stronger-multi-scale-deep-graph-convolutional-networks>
AUTHORS: Sitao Luan, Mingde Zhao, Xiao-Wen Chang, Doina Precup
HIGHLIGHT: In this paper, we first analyze key factors constraining the expressive power of existing Graph Convolutional Networks (GCNs), including the activation function and shallow learning mechanisms. Then, we generalize spectral graph convolution and deep GCN in block Krylov subspace forms, upon which we devise two architectures, both scalable in depth however making use of multi-scale information differently.
- 982, TITLE: Statistical Model Aggregation via Parameter Matching
<https://papers.nips.cc/paper/9277-statistical-model-aggregation-via-parameter-matching>
AUTHORS: Mikhail Yurochkin, Mayank Agarwal, Soumya Ghosh, Kristjan Greenewald, Nghia Hoang
HIGHLIGHT: Exploiting tools from Bayesian nonparametrics, we develop a general meta-modeling framework that learns shared global latent structures by identifying correspondences among local model parameterizations.
- 983, TITLE: On the (In)fidelity and Sensitivity of Explanations
<https://papers.nips.cc/paper/9278-on-the-infidelity-and-sensitivity-of-explanations>
AUTHORS: Chih-Kuan Yeh, Cheng-Yu Hsieh, Arun Suggala, David I. Inouye, Pradeep K. Ravikumar
HIGHLIGHT: We propose simple robust variants of two notions that have been considered in recent literature: (in)fidelity, and sensitivity.
- 984, TITLE: Exponential Family Estimation via Adversarial Dynamics Embedding
<https://papers.nips.cc/paper/9279-exponential-family-estimation-via-adversarial-dynamics-embedding>
AUTHORS: Bo Dai, Zhen Liu, Hanjun Dai, Niao He, Arthur Gretton, Le Song, Dale Schuurmans
HIGHLIGHT: We present an efficient algorithm for maximum likelihood estimation (MLE) of exponential family models, with a general parametrization of the energy function that includes neural networks.
- 985, TITLE: The Broad Optimality of Profile Maximum Likelihood
<https://papers.nips.cc/paper/9280-the-broad-optimality-of-profile-maximum-likelihood>
AUTHORS: Yi Hao, Alon Orlitsky
HIGHLIGHT: We study three fundamental statistical-learning problems: distribution estimation, property estimation, and property testing.
- 986, TITLE: MintNet: Building Invertible Neural Networks with Masked Convolutions
<https://papers.nips.cc/paper/9281-mintnet-building-invertible-neural-networks-with-masked-convolutions>
AUTHORS: Yang Song, Chenlin Meng, Stefano Ermon
HIGHLIGHT: We propose a new way of constructing invertible neural networks by combining simple building blocks with a novel set of composition rules.
- 987, TITLE: Information-Theoretic Generalization Bounds for SGLD via Data-Dependent Estimates
<https://papers.nips.cc/paper/9282-information-theoretic-generalization-bounds-for-sgld-via-data-dependent-estimates>
AUTHORS: Jeffrey Negrea, Mahdi Haghifam, Gintare Karolina Dziugaite, Ashish Khisti, Daniel M. Roy
HIGHLIGHT: In this work, we improve upon the stepwise analysis of noisy iterative learning algorithms initiated by Pensia, Jog, and Loh (2018) and recently extended by Bu, Zou, and Veeravalli (2019).
- 988, TITLE: On Distributed Averaging for Stochastic k-PCA
<https://papers.nips.cc/paper/9283-on-distributed-averaging-for-stochastic-k-pca>
AUTHORS: Aditya Bhaskara, Pruthuvi Maheshakya Wijewardena

HIGHLIGHT: We consider a slight variant of the well-studied "distributed averaging" approach, and prove that this leads to significantly better bounds on the dependence between 'n' and the eigenvalue gaps.

989, **TITLE:** Controllable Unsupervised Text Attribute Transfer via Editing Entangled Latent Representation
<https://papers.nips.cc/paper/9284-controllable-unsupervised-text-attribute-transfer-via-editing-entangled-latent-representation>
AUTHORS: Ke Wang, Hang Hua, Xiaojun Wan
HIGHLIGHT: To address the above problems, we propose a more flexible unsupervised text attribute transfer framework which replaces the process of modeling attribute with minimal editing of latent representations based on an attribute classifier.

990, **TITLE:** MaxGap Bandit: Adaptive Algorithms for Approximate Ranking
<https://papers.nips.cc/paper/9285-maxgap-bandit-adaptive-algorithms-for-approximate-ranking>
AUTHORS: Sumeet Katariya, Ardhendu Tripathy, Robert Nowak
HIGHLIGHT: We propose elimination and UCB-style algorithms and show that they are minimax optimal.

991, **TITLE:** Bias Correction of Learned Generative Models using Likelihood-Free Importance Weighting
<https://papers.nips.cc/paper/9286-bias-correction-of-learned-generative-models-using-likelihood-free-importance-weighting>
AUTHORS: Aditya Grover, Jiaming Song, Ashish Kapoor, Kenneth Tran, Alekh Agarwal, Eric J. Horvitz, Stefano Ermon
HIGHLIGHT: We employ this likelihood-free importance weighting method to correct for the bias in generative models.

992, **TITLE:** Online Forecasting of Total-Variation-bounded Sequences
<https://papers.nips.cc/paper/9287-online-forecasting-of-total-variation-bounded-sequences>
AUTHORS: Dheeraj Baby, Yu-Xiang Wang
HIGHLIGHT: We consider the problem of online forecasting of sequences of length n with total-variation at most C_n using observations contaminated by independent σ -subgaussian noise.

993, **TITLE:** Local SGD with Periodic Averaging: Tighter Analysis and Adaptive Synchronization
<https://papers.nips.cc/paper/9288-local-sgd-with-periodic-averaging-tighter-analysis-and-adaptive-synchronization>
AUTHORS: Farzin Haddadpour, Mohammad Mahdi Kamani, Mehrdad Mahdavi, Viveck Cadambe
HIGHLIGHT: In this paper, we study local distributed SGD, where data is partitioned among computation nodes, and the computation nodes perform local updates with periodically exchanging the model among the workers to perform averaging.

994, **TITLE:** Data Parameters: A New Family of Parameters for Learning a Differentiable Curriculum
<https://papers.nips.cc/paper/9289-data-parameters-a-new-family-of-parameters-for-learning-a-differentiable-curriculum>
AUTHORS: Shreyas Saxena, Oncel Tuzel, Dennis DeCoste
HIGHLIGHT: In this work, we address this problem by introducing data parameters.

995, **TITLE:** Unified Sample-Optimal Property Estimation in Near-Linear Time
<https://papers.nips.cc/paper/9290-unified-sample-optimal-property-estimation-in-near-linear-time>
AUTHORS: Yi Hao, Alon Orlitsky
HIGHLIGHT: We consider the fundamental learning problem of estimating properties of distributions over large domains.

996, **TITLE:** Region Mutual Information Loss for Semantic Segmentation
<https://papers.nips.cc/paper/9291-region-mutual-information-loss-for-semantic-segmentation>
AUTHORS: Shuai Zhao, Yang Wang, Zheng Yang, Deng Cai
HIGHLIGHT: In this paper, we develop a region mutual information (RMI) loss to model the dependencies among pixels more simply and efficiently.

997, **TITLE:** Learning Stable Deep Dynamics Models
<https://papers.nips.cc/paper/9292-learning-stable-deep-dynamics-models>
AUTHORS: J. Zico Kolter, Gaurav Manek
HIGHLIGHT: In this paper, we propose an approach for learning dynamical systems that are guaranteed to be stable over the entire state space.

998, **TITLE:** Image Captioning: Transforming Objects into Words
<https://papers.nips.cc/paper/9293-image-captioning-transforming-objects-into-words>
AUTHORS: Simao Herdade, Armin Kappeler, Kofi Boakye, Joao Soares
HIGHLIGHT: In this work we introduce the Object Relation Transformer, that builds upon this approach by explicitly incorporating information about the spatial relationship between input detected objects through geometric attention.

- 999, TITLE: Greedy Sampling for Approximate Clustering in the Presence of Outliers
<https://papers.nips.cc/paper/9294-greedy-sampling-for-approximate-clustering-in-the-presence-of-outliers>
AUTHORS: Aditya Bhaskara, Sharvaree Vadgama, Hong Xu
HIGHLIGHT: In this work we show that for k-means and k-center clustering, simple modifications to the well-studied greedy algorithms result in nearly identical guarantees, while additionally being robust to outliers.
- 1000, TITLE: Adversarial Fisher Vectors for Unsupervised Representation Learning
<https://papers.nips.cc/paper/9295-adversarial-fisher-vectors-for-unsupervised-representation-learning>
AUTHORS: Joshua Susskind, Shuangfei Zhai, Walter Talbott, Carlos Guestrin
HIGHLIGHT: We examine Generative Adversarial Networks (GANs) through the lens of deep Energy Based Models (EBMs), with the goal of exploiting the density model that follows from this formulation.
- 1001, TITLE: On Tractable Computation of Expected Predictions
<https://papers.nips.cc/paper/9296-on-tractable-computation-of-expected-predictions>
AUTHORS: Pasha Khosravi, YooJung Choi, Yitao Liang, Antonio Vergari, Guy Van den Broeck
HIGHLIGHT: In this paper, we identify a pair of generative and discriminative models that enables tractable computation of expectations, as well as moments of any order, of the latter with respect to the former in case of regression.
- 1002, TITLE: Levenshtein Transformer
<https://papers.nips.cc/paper/9297-levenshtein-transformer>
AUTHORS: Jiatao Gu, Changhan Wang, Junbo Zhao
HIGHLIGHT: In this work, we develop Levenshtein Transformer, a new partially autoregressive model devised for more flexible and amenable sequence generation.
- 1003, TITLE: Unlabeled Data Improves Adversarial Robustness
<https://papers.nips.cc/paper/9298-unlabeled-data-improves-adversarial-robustness>
AUTHORS: Yair Carmon, Aditi Raghunathan, Ludwig Schmidt, John C. Duchi, Percy S. Liang
HIGHLIGHT: We prove that unlabeled data bridges this gap: a simple semisupervised learning procedure (self-training) achieves high robust accuracy using the same number of labels required for achieving high standard accuracy.
- 1004, TITLE: Machine Teaching of Active Sequential Learners
<https://papers.nips.cc/paper/9299-machine-teaching-of-active-sequential-learners>
AUTHORS: Tomi Peltola, Mustafa Mert ?elikok, Pedram Daei, Samuel Kaski
HIGHLIGHT: We formulate this sequential teaching problem, which current techniques in machine teaching do not address, as a Markov decision process, with the dynamics nesting a model of the learner and the actions being the teacher's responses.
- 1005, TITLE: Gaussian-Based Pooling for Convolutional Neural Networks
<https://papers.nips.cc/paper/9300-gaussian-based-pooling-for-convolutional-neural-networks>
AUTHORS: Takumi Kobayashi
HIGHLIGHT: In this paper, to improve performance of CNNs, we propose a novel local pooling method based on the Gaussian-based probabilistic model over local neuron activations for flexibly pooling (extracting) features, in contrast to the previous model restricting the output within the convex hull of local neurons.
- 1006, TITLE: Meta Architecture Search
<https://papers.nips.cc/paper/9301-meta-architecture-search>
AUTHORS: Albert Shaw, Wei Wei, Weiyang Liu, Le Song, Bo Dai
HIGHLIGHT: We propose the Bayesian Meta Architecture SEArch (BASE) framework which takes advantage of a Bayesian formulation of the architecture search problem to learn over an entire set of tasks simultaneously.
- 1007, TITLE: NAOMI: Non-Autoregressive Multiresolution Sequence Imputation
<https://papers.nips.cc/paper/9302-naomi-non-autoregressive-multiresolution-sequence-imputation>
AUTHORS: Yukai Liu, Rose Yu, Stephan Zheng, Eric Zhan, Yisong Yue
HIGHLIGHT: In this paper, we take a non-autoregressive approach and propose a novel deep generative model: Non-Autoregressive Multiresolution Imputation (NAOMI) to impute long-range sequences given arbitrary missing patterns.
- 1008, TITLE: Layer-Dependent Importance Sampling for Training Deep and Large Graph Convolutional Networks
<https://papers.nips.cc/paper/9303-layer-dependent-importance-sampling-for-training-deep-and-large-graph-convolutional-networks>
AUTHORS: Difan Zou, Ziniu Hu, Yewen Wang, Song Jiang, Yizhou Sun, Quanquan Gu
HIGHLIGHT: To deal with the above two problems, we propose a new effective sampling algorithm called LAyer-Dependent ImportancE Sampling (LADIES).

1009, TITLE: Two Generator Game: Learning to Sample via Linear Goodness-of-Fit Test
<https://papers.nips.cc/paper/9304-two-generator-game-learning-to-sample-via-linear-goodness-of-fit-test>
AUTHORS: Lizhong Ding, Mengyang Yu, Li Liu, Fan Zhu, Yong Liu, Yu Li, Ling Shao
HIGHLIGHT: To solve this problem, we formulate a deep energy adversarial network (DEAN), which casts the energy model learned from real data into an optimization of a goodness-of-fit (GOF) test statistic.

1010, TITLE: Distribution oblivious, risk-aware algorithms for multi-armed bandits with unbounded rewards
<https://papers.nips.cc/paper/9305-distribution-oblivious-risk-aware-algorithms-for-multi-armed-bandits-with-unbounded-rewards>
AUTHORS: Anmol Kaglecha, Jayakrishnan Nair, Krishna Jagannathan
HIGHLIGHT: In this paper, we consider the problem of selecting the arm that optimizes a linear combination of the expected reward and the associated Conditional Value at Risk (CVaR) in a fixed budget best-arm identification framework.

1011, TITLE: Private Stochastic Convex Optimization with Optimal Rates
<https://papers.nips.cc/paper/9306-private-stochastic-convex-optimization-with-optimal-rates>
AUTHORS: Raef Bassily, Vitaly Feldman, Kunal Talwar, Abhradeep Guha Thakurta
HIGHLIGHT: We study differentially private (DP) algorithms for stochastic convex optimization (SCO).

1012, TITLE: Provably Robust Deep Learning via Adversarially Trained Smoothed Classifiers
<https://papers.nips.cc/paper/9307-provably-robust-deep-learning-via-adversarially-trained-smoothed-classifiers>
AUTHORS: Hadi Salman, Jerry Li, Ilya Razenshteyn, Pengchuan Zhang, Huan Zhang, Sebastien Bubeck, Greg Yang
HIGHLIGHT: In this paper, we employ adversarial training to improve the performance of randomized smoothing.

1013, TITLE: Demystifying Black-box Models with Symbolic Metamodels
<https://papers.nips.cc/paper/9308-demystifying-black-box-models-with-symbolic-metamodels>
AUTHORS: Ahmed M. Alaa, Mihaela van der Schaar
HIGHLIGHT: To address this issue, we introduce the symbolic metamodeling framework - a general methodology for interpreting predictions by converting "black-box" models into "white-box" functions that are understandable to human subjects.

1014, TITLE: Neural Temporal-Difference Learning Converges to Global Optima
<https://papers.nips.cc/paper/9309-neural-temporal-difference-learning-converges-to-global-optima>
AUTHORS: Qi Cai, Zhuoran Yang, Jason D. Lee, Zhaoran Wang
HIGHLIGHT: In this paper, we prove for the first time that neural TD converges at a sublinear rate to the global optimum of the mean-squared projected Bellman error for policy evaluation.

1015, TITLE: Privacy-Preserving Q-Learning with Functional Noise in Continuous Spaces
<https://papers.nips.cc/paper/9310-privacy-preserving-q-learning-with-functional-noise-in-continuous-spaces>
AUTHORS: Baoxiang Wang, Nidhi Hegde
HIGHLIGHT: Our aim is to protect the value function approximator, without regard to the number of states queried to the function.

1016, TITLE: Attentive State-Space Modeling of Disease Progression
<https://papers.nips.cc/paper/9311-attentive-state-space-modeling-of-disease-progression>
AUTHORS: Ahmed M. Alaa, Mihaela van der Schaar
HIGHLIGHT: In this paper, we develop the attentive state-space model, a deep probabilistic model that learns accurate and interpretable structured representations for disease trajectories.

1017, TITLE: Online EXP3 Learning in Adversarial Bandits with Delayed Feedback
<https://papers.nips.cc/paper/9312-online-exp3-learning-in-adversarial-bandits-with-delayed-feedback>
AUTHORS: Ilai Bistritz, Zhengyuan Zhou, Xi Chen, Nicholas Bambos, Jose Blanchet
HIGHLIGHT: For the case where $\sum_{t=1}^T d_t$ and T are unknown, we propose a novel doubling trick for online learning with delays and prove that this adaptive EXP3 achieves a regret of $O(\sqrt{\ln K \left(K^2 T + \sum_{t=1}^T d_t \right)})$.

1018, TITLE: A Direct $\tilde{O}(1/\epsilon)$ Iteration Parallel Algorithm for Optimal Transport
[https://papers.nips.cc/paper/9313-a-direct-tilde{O}\(1/epsilon\)-iteration-parallel-algorithm-for-optimal-transport](https://papers.nips.cc/paper/9313-a-direct-tilde{O}(1/epsilon)-iteration-parallel-algorithm-for-optimal-transport)
AUTHORS: Arun Jambulapati, Aaron Sidford, Kevin Tian
HIGHLIGHT: We give an algorithm which solves the problem to additive ϵ accuracy with $\tilde{O}(1/\epsilon)$ parallel depth and $\tilde{O}(n^2/\epsilon)$ work.

- 1019, TITLE: Faster Boosting with Smaller Memory
<https://papers.nips.cc/paper/9314-faster-boosting-with-smaller-memory>
AUTHORS: Julaiti Alafate, Yoav S. Freund
HIGHLIGHT: This paper presents an alternative approach to implementing the boosted trees, which achieves a significant speedup over XGBoost and LightGBM, especially when the memory size is small.
- 1020, TITLE: Variance Reduction for Matrix Games
<https://papers.nips.cc/paper/9315-variance-reduction-for-matrix-games>
AUTHORS: Yair Carmon, Yujia Jin, Aaron Sidford, Kevin Tian
HIGHLIGHT: We present a randomized primal-dual algorithm that solves the problem $\min_x \max_y y^T A x$ to additive error ϵ in time $\text{nnz}(A) + \sqrt{\text{nnz}(A)} n / \epsilon$, for matrix A with larger dimension n and $\text{nnz}(A)$ nonzero entries.
- 1021, TITLE: Learning Neural Networks with Adaptive Regularization
<https://papers.nips.cc/paper/9316-learning-neural-networks-with-adaptive-regularization>
AUTHORS: Han Zhao, Yao-Hung Hubert Tsai, Russ R. Salakhutdinov, Geoffrey J. Gordon
HIGHLIGHT: To optimize the model, we present an efficient block coordinate descent algorithm with analytical solutions.
- 1022, TITLE: Distributed estimation of the inverse Hessian by determinantal averaging
<https://papers.nips.cc/paper/9317-distributed-estimation-of-the-inverse-hessian-by-determinantal-averaging>
AUTHORS: Michal Dereziński, Michael W. Mahoney
HIGHLIGHT: To address this, we propose determinantal averaging, a new approach for correcting the inversion bias.
- 1023, TITLE: Smoothing Structured Decomposable Circuits
<https://papers.nips.cc/paper/9318-smoothing-structured-decomposable-circuits>
AUTHORS: Andy Shih, Guy Van den Broeck, Paul Beame, Antoine Amarilli
HIGHLIGHT: We propose a near-linear time algorithm for this task and explore lower bounds for smoothing decomposable circuits, using existing results on range-sum queries.
- 1024, TITLE: Efficient and Accurate Estimation of Lipschitz Constants for Deep Neural Networks
<https://papers.nips.cc/paper/9319-efficient-and-accurate-estimation-of-lipschitz-constants-for-deep-neural-networks>
AUTHORS: Mahyar Fazlyab, Alexander Robey, Hamed Hassani, Manfred Morari, George Pappas
HIGHLIGHT: In this paper, we present a convex optimization framework to compute guaranteed upper bounds on the Lipschitz constant of DNNs both accurately and efficiently.
- 1025, TITLE: Provable Non-linear Inductive Matrix Completion
<https://papers.nips.cc/paper/9320-provable-non-linear-inductive-matrix-completion>
AUTHORS: Kai Zhong, Zhao Song, Prateek Jain, Inderjit S. Dhillon
HIGHLIGHT: In this paper, we provide the first theoretical analysis for a simple NIMC model in the realizable setting, where the relevance score of a (query, item) pair is formulated as the inner product between their single-layer neural representations.
- 1026, TITLE: Communication-Efficient Distributed Blockwise Momentum SGD with Error-Feedback
<https://papers.nips.cc/paper/9321-communication-efficient-distributed-blockwise-momentum-sgd-with-error-feedback>
AUTHORS: Shuai Zheng, Ziyue Huang, James Kwok
HIGHLIGHT: In this paper, we propose a general distributed compressed SGD with Nesterov's momentum.
- 1027, TITLE: Sparse Variational Inference: Bayesian Coresets from Scratch
<https://papers.nips.cc/paper/9322-sparse-variational-inference-bayesian-coresets-from-scratch>
AUTHORS: Trevor Campbell, Boyan Beronov
HIGHLIGHT: In the present work we remove this requirement by formulating coreset construction as sparsity-constrained variational inference within an exponential family.
- 1028, TITLE: Personalizing Many Decisions with High-Dimensional Covariates
<https://papers.nips.cc/paper/9323-personalizing-many-decisions-with-high-dimensional-covariates>
AUTHORS: Nima Hamidi, Mohsen Bayati, Kapil Gupta
HIGHLIGHT: The main contribution of this paper is to introduce and theoretically analyze a new algorithm (REAL Bandit) with a regret that scales by $r^2(k+d)$ when r is rank of the k by d matrix of unknown parameters.
- 1029, TITLE: A Necessary and Sufficient Stability Notion for Adaptive Generalization
<https://papers.nips.cc/paper/9324-a-necessary-and-sufficient-stability-notion-for-adaptive-generalization>
AUTHORS: Moshe Shenfeld, Katrina Ligett

HIGHLIGHT: We introduce a new notion of the stability of computations, which holds under post-processing and adaptive composition.

1030, TITLE: Necessary and Sufficient Geometries for Gradient Methods
<https://papers.nips.cc/paper/9325-necessary-and-sufficient-geometries-for-gradient-methods>

AUTHORS: Daniel Levy, John C. Duchi

HIGHLIGHT: We study the impact of the constraint set and gradient geometry on the convergence of online and stochastic methods for convex optimization, providing a characterization of the geometries for which stochastic gradient and adaptive gradient methods are (minimax) optimal.

1031, TITLE: Landmark Ordinal Embedding
<https://papers.nips.cc/paper/9326-landmark-ordinal-embedding>

AUTHORS: Nikhil Ghosh, Yuxin Chen, Yisong Yue

HIGHLIGHT: In this paper, we aim to learn a low-dimensional Euclidean representation from a set of constraints of the form "item j is closer to item i than item k".

1032, TITLE: Identification of Conditional Causal Effects under Markov Equivalence
<https://papers.nips.cc/paper/9327-identification-of-conditional-causal-effects-under-markov-equivalence>

AUTHORS: Amin Jaber, Jiji Zhang, Elias Bareinboim

HIGHLIGHT: In this work, we derive an algorithm to identify conditional effects, which are particularly useful for evaluating conditional plans or policies.

1033, TITLE: The Thermodynamic Variational Objective
<https://papers.nips.cc/paper/9328-the-thermodynamic-variational-objective>

AUTHORS: Vaden Masrani, Tuan Anh Le, Frank Wood

HIGHLIGHT: We introduce the thermodynamic variational objective (TVO) for learning in both continuous and discrete deep generative models.

1034, TITLE: Global Guarantees for Blind Demodulation with Generative Priors
<https://papers.nips.cc/paper/9329-global-guarantees-for-blind-demodulation-with-generative-priors>

AUTHORS: Paul Hand, Babhru Joshi

HIGHLIGHT: We study a deep learning inspired formulation for the blind demodulation problem, which is the task of recovering two unknown vectors from their entrywise multiplication.

1035, TITLE: Exact sampling of determinantal point processes with sublinear time preprocessing
<https://papers.nips.cc/paper/9330-exact-sampling-of-determinantal-point-processes-with-sublinear-time-preprocessing>

AUTHORS: Michal Dereziński, Daniele Calandriello, Michal Valko

HIGHLIGHT: For this purpose we provide DPP-VFX, a new algorithm which, given access only to L , samples exactly from a determinantal point process while satisfying the following two properties: (1) its preprocessing cost is $n \text{ poly}(k)$, i.e., sublinear in the size of L , and (2) its sampling cost is $\text{poly}(k)$, i.e., independent of the size of L . Prior to our results, state-of-the-art exact samplers required $O(n^3)$ preprocessing time and sampling time linear in n or dependent on the spectral properties of L .

1036, TITLE: Geometry-Aware Neural Rendering
<https://papers.nips.cc/paper/9331-geometry-aware-neural-rendering>

AUTHORS: Joshua Tobin, Wojciech Zaremba, Pieter Abbeel

HIGHLIGHT: We propose Epipolar Cross Attention (ECA), an attention mechanism that leverages the geometry of the scene to perform efficient non-local operations, requiring only $O(n)$ comparisons per spatial dimension instead of $O(n^2)$.

1037, TITLE: Variational Temporal Abstraction
<https://papers.nips.cc/paper/9332-variational-temporal-abstraction>

AUTHORS: Taesup Kim, Sungjin Ahn, Yoshua Bengio

HIGHLIGHT: We introduce a variational approach to learning and inference of temporally hierarchical structure and representation for sequential data.

1038, TITLE: Subquadratic High-Dimensional Hierarchical Clustering
<https://papers.nips.cc/paper/9333-subquadratic-high-dimensional-hierarchical-clustering>

AUTHORS: Amir Abboud, Vincent Cohen-Addad, Hussein Houdrouge

HIGHLIGHT: We consider the widely-used average-linkage, single-linkage, and Ward's methods for computing hierarchical clusterings of high-dimensional Euclidean inputs.

- 1039, TITLE: Learning Auctions with Robust Incentive Guarantees
<https://papers.nips.cc/paper/9334-learning-auctions-with-robust-incentive-guarantees>
AUTHORS: Jacob D. Abernethy, Rachel Cummings, Bhuvesh Kumar, Sam Taggart, Jamie H. Morgenstern
HIGHLIGHT: In this paper, we combine tools from differential privacy, mechanism design, and sample complexity to give a repeated auction that (1) learns bidder demand from past data, (2) is approximately revenue-optimal, and (3) strategically robust, as it incentivizes bidders to behave truthfully.
- 1040, TITLE: Policy Optimization Provably Converges to Nash Equilibria in Zero-Sum Linear Quadratic Games
<https://papers.nips.cc/paper/9335-policy-optimization-provably-converges-to-nash-equilibria-in-zero-sum-linear-quadratic-games>
AUTHORS: Kaiqing Zhang, Zhuoran Yang, Tamer Basar
HIGHLIGHT: Building upon this, we develop three projected nested-gradient methods that are guaranteed to converge to the NE of the game.
- 1041, TITLE: Uniform convergence may be unable to explain generalization in deep learning
<https://papers.nips.cc/paper/9336-uniform-convergence-may-be-unable-to-explain-generalization-in-deep-learning>
AUTHORS: Vaishnavh Nagarajan, J. Zico Kolter
HIGHLIGHT: Through these findings, we cast doubt on the power of uniform convergence-based generalization bounds to provide a complete picture of why overparameterized deep networks generalize well.
- 1042, TITLE: A Zero-Positive Learning Approach for Diagnosing Software Performance Regressions
<https://papers.nips.cc/paper/9337-a-zero-positive-learning-approach-for-diagnosing-software-performance-regressions>
AUTHORS: Mejbah Alam, Justin Gottschlich, Nesime Tatbul, Javier S. Turek, Tim Mattson, Abdullah Muzahid
HIGHLIGHT: In this paper, we apply MP to the automation of software performance regression testing.
- 1043, TITLE: DTWNet: a Dynamic Time Warping Network
<https://papers.nips.cc/paper/9338-dtwnet-a-dynamic-time-warping-network>
AUTHORS: Xingyu Cai, Tingyang Xu, Jinfeng Yi, Junzhou Huang, Sanguthevar Rajasekaran
HIGHLIGHT: In this paper, we propose a novel component in an artificial neural network.
- 1044, TITLE: Structured Graph Learning Via Laplacian Spectral Constraints
<https://papers.nips.cc/paper/9339-structured-graph-learning-via-laplacian-spectral-constraints>
AUTHORS: Sandeep Kumar, Jiayi Ying, Jose Vinicius de Miranda Cardoso, Daniel Palomar
HIGHLIGHT: In this paper, we first show, for a set of important graph families it is possible to convert the combinatorial constraints of structure into eigenvalue constraints of the graph Laplacian matrix. Then we introduce a unified graph learning framework lying at the integration of the spectral properties of the Laplacian matrix with Gaussian graphical modeling, which is capable of learning structures of a large class of graph families.
- 1045, TITLE: Thresholding Bandit with Optimal Aggregate Regret
<https://papers.nips.cc/paper/9340-thresholding-bandit-with-optimal-aggregate-regret>
AUTHORS: Chao Tao, Sa'ul Blanco, Jian Peng, Yuan Zhou
HIGHLIGHT: We introduce LSA, a new, simple and anytime algorithm that aims to minimize the aggregate regret (or the expected number of mis-classified arms).
- 1046, TITLE: Towards Explaining the Regularization Effect of Initial Large Learning Rate in Training Neural Networks
<https://papers.nips.cc/paper/9341-towards-explaining-the-regularization-effect-of-initial-large-learning-rate-in-training-neural-networks>
AUTHORS: Yuanzhi Li, Colin Wei, Tengyu Ma
HIGHLIGHT: Towards explaining this phenomenon, we devise a setting in which we can prove that a two layer network trained with large initial learning rate and annealing provably generalizes better than the same network trained with a small learning rate from the start.
- 1047, TITLE: Rethinking Kernel Methods for Node Representation Learning on Graphs
<https://papers.nips.cc/paper/9342-rethinking-kernel-methods-for-node-representation-learning-on-graphs>
AUTHORS: Yu Tian, Long Zhao, Xi Peng, Dimitris Metaxas
HIGHLIGHT: Here, we present a novel theoretical kernel-based framework for node classification that can bridge the gap between these two representation learning problems on graphs.
- 1048, TITLE: Causal Confusion in Imitation Learning
<https://papers.nips.cc/paper/9343-causal-confusion-in-imitation-learning>
AUTHORS: Pim de Haan, Dinesh Jayaraman, Sergey Levine

HIGHLIGHT: We investigate how this problem arises, and propose a solution to combat it through targeted interventions---either environment interaction or expert queries---to determine the correct causal model.

1049, TITLE: Optimizing Generalized PageRank Methods for Seed-Expansion Community Detection
<https://papers.nips.cc/paper/9344-optimizing-generalized-pagerank-methods-for-seed-expansion-community-detection>
AUTHORS: Pan Li, I Chien, Olgica Milenkovic
HIGHLIGHT: Given this result, we propose a new GPR, termed Inverse PR (IPR), with LP weights that increase for the initial few steps of the walks.

1050, TITLE: The Case for Evaluating Causal Models Using Interventional Measures and Empirical Data
<https://papers.nips.cc/paper/9345-the-case-for-evaluating-causal-models-using-interventional-measures-and-empirical-data>
AUTHORS: Amanda Gentzel, Dan Garant, David Jensen
HIGHLIGHT: We argue for more frequent use of evaluation techniques that examine interventional measures rather than structural or observational measures, and that evaluate those measures on empirical data rather than synthetic data.

1051, TITLE: Dimension-Free Bounds for Low-Precision Training
<https://papers.nips.cc/paper/9346-dimension-free-bounds-for-low-precision-training>
AUTHORS: Zheng Li, Christopher M. De Sa
HIGHLIGHT: In this paper, we derive new bounds for low-precision training algorithms that do not contain the dimension d , which lets us better understand what affects the convergence of these algorithms as parameters scale.

1052, TITLE: Concentration of risk measures: A Wasserstein distance approach
<https://papers.nips.cc/paper/9347-concentration-of-risk-measures-a-wasserstein-distance-approach>
AUTHORS: Sanjay P. Bhat, Prashanth L.A.
HIGHLIGHT: Concentration of risk measures: A Wasserstein distance approach

1053, TITLE: Meta-Inverse Reinforcement Learning with Probabilistic Context Variables
<https://papers.nips.cc/paper/9348-meta-inverse-reinforcement-learning-with-probabilistic-context-variables>
AUTHORS: Lantao Yu, Tianhe Yu, Chelsea Finn, Stefano Ermon
HIGHLIGHT: To this end, we propose a deep latent variable model that is capable of learning rewards from unstructured, multi-task demonstration data, and critically, use this experience to infer robust rewards for new, structurally-similar tasks from a single demonstration.

1054, TITLE: Stabilizing Off-Policy Q-Learning via Bootstrapping Error Reduction
<https://papers.nips.cc/paper/9349-stabilizing-off-policy-q-learning-via-bootstrapping-error-reduction>
AUTHORS: Aviral Kumar, Justin Fu, Matthew Soh, George Tucker, Sergey Levine
HIGHLIGHT: Based on our analysis, we propose a practical algorithm, bootstrapping error accumulation reduction (BEAR).

1055, TITLE: Bayesian Optimization with Unknown Search Space
<https://papers.nips.cc/paper/9350-bayesian-optimization-with-unknown-search-space>
AUTHORS: Huong Ha, Santu Rana, Sunil Gupta, Thanh Nguyen, Hung Tran-The, Svetha Venkatesh
HIGHLIGHT: To address this problem, we propose a systematic volume expansion strategy for the Bayesian optimization.

1056, TITLE: On the Downstream Performance of Compressed Word Embeddings
<https://papers.nips.cc/paper/9351-on-the-downstream-performance-of-compressed-word-embeddings>
AUTHORS: Avner May, Jian Zhang, Tri Dao, Christopher R?
HIGHLIGHT: We thus propose the eigenspace overlap score as a new measure.

1057, TITLE: Multivariate Distributionally Robust Convex Regression under Absolute Error Loss
<https://papers.nips.cc/paper/9352-multivariate-distributionally-robust-convex-regression-under-absolute-error-loss>
AUTHORS: Jose Blanchet, Peter W. Glynn, Jun Yan, Zhengqing Zhou
HIGHLIGHT: This paper proposes a novel non-parametric multidimensional convex regression estimator which is designed to be robust to adversarial perturbations in the empirical measure.

1058, TITLE: Neural Relational Inference with Fast Modular Meta-learning
<https://papers.nips.cc/paper/9353-neural-relational-inference-with-fast-modular-meta-learning>
AUTHORS: Ferran Alet, Erica Weng, Tom?s Lozano-Perez, Leslie Pack Kaelbling
HIGHLIGHT: We frame relational inference as a modular meta-learning problem, where neural modules are trained to be composed in different ways to solve many tasks.

- 1059, TITLE: Gradient based sample selection for online continual learning
<https://papers.nips.cc/paper/9354-gradient-based-sample-selection-for-online-continual-learning>
AUTHORS: Rahaf Aljundi, Min Lin, Baptiste Goujaud, Yoshua Bengio
HIGHLIGHT: In this work, we formulate sample selection as a constraint reduction problem based on the constrained optimization view of continual learning.
- 1060, TITLE: Attribution-Based Confidence Metric For Deep Neural Networks
<https://papers.nips.cc/paper/9355-attribution-based-confidence-metric-for-deep-neural-networks>
AUTHORS: Susmit Jha, Sunny Raj, Steven Fernandes, Sumit K. Jha, Somesh Jha, Brian Jalaian, Gunjan Verma, Ananthram Swami
HIGHLIGHT: We propose a novel confidence metric, namely, attribution-based confidence (ABC) for deep neural networks (DNNs).
- 1061, TITLE: Theoretical evidence for adversarial robustness through randomization
<https://papers.nips.cc/paper/9356-theoretical-evidence-for-adversarial-robustness-through-randomization>
AUTHORS: Rafael Pinot, Laurent Meunier, Alexandre Araujo, Hisashi Kashima, Florian Yger, Cedric Gouy-Pailler, Jamal Atif
HIGHLIGHT: This paper investigates the theory of robustness against adversarial attacks.
- 1062, TITLE: Online Continual Learning with Maximal Interfered Retrieval
<https://papers.nips.cc/paper/9357-online-continual-learning-with-maximal-interfered-retrieval>
AUTHORS: Rahaf Aljundi, Eugene Belilovsky, Tinne Tuytelaars, Laurent Charlin, Massimo Caccia, Min Lin, Lucas Page-Caccia
HIGHLIGHT: In this work, we consider a controlled sampling of memories for replay.
- 1063, TITLE: Neural Attribution for Semantic Bug-Localization in Student Programs
<https://papers.nips.cc/paper/9358-neural-attribution-for-semantic-bug-localization-in-student-programs>
AUTHORS: Rahul Gupta, Aditya Kanade, Shirish Shevade
HIGHLIGHT: In this work, we present NeuralBugLocator, a deep learning based technique, that can localize the bugs in a faulty program with respect to a failing test, without even running the program.
- 1064, TITLE: Adaptive Temporal-Difference Learning for Policy Evaluation with Per-State Uncertainty Estimates
<https://papers.nips.cc/paper/9359-adaptive-temporal-difference-learning-for-policy-evaluation-with-per-state-uncertainty-estimates>
AUTHORS: Carlos Riquelme, Hugo Penedones, Damien Vincent, Hartmut Maennel, Sylvain Gelly, Timothy A. Mann, Andre Barreto, Gergely Neu
HIGHLIGHT: In this paper, we argue that the larger bias of TD can be a result of the amplification of local approximation errors.
- 1065, TITLE: SPoC: Search-based Pseudocode to Code
<https://papers.nips.cc/paper/9360-spoc-search-based-pseudocode-to-code>
AUTHORS: Sumith Kulal, Panupong Pasupat, Kartik Chandra, Mina Lee, Oded Padon, Alex Aiken, Percy S. Liang
HIGHLIGHT: We consider the task of mapping pseudocode to executable code, assuming a one-to-one correspondence between lines of pseudocode and lines of code.
- 1066, TITLE: Generative Modeling by Estimating Gradients of the Data Distribution
<https://papers.nips.cc/paper/9361-generative-modeling-by-estimating-gradients-of-the-data-distribution>
AUTHORS: Yang Song, Stefano Ermon
HIGHLIGHT: We introduce a new generative model where samples are produced via Langevin dynamics using gradients of the data distribution estimated with score matching.
- 1067, TITLE: Adversarial Music: Real world Audio Adversary against Wake-word Detection System
<https://papers.nips.cc/paper/9362-adversarial-music-real-world-audio-adversary-against-wake-word-detection-system>
AUTHORS: Juncheng Li, Shuhui Qu, Xinjian Li, Joseph Szurley, J. Zico Kolter, Florian Metze
HIGHLIGHT: In this work, we target our attack on the wake-word detection system.
- 1068, TITLE: Prediction of Spatial Point Processes: Regularized Method with Out-of-Sample Guarantees
<https://papers.nips.cc/paper/9363-prediction-of-spatial-point-processes-regularized-method-with-out-of-sample-guarantees>
AUTHORS: Muhammad Osama, Dave Zachariah, Peter Stoica
HIGHLIGHT: In this paper, we develop a method to infer predictive intensity intervals by learning a spatial model using a regularized criterion.

- 1069, TITLE: Debiased Bayesian inference for average treatment effects
<https://papers.nips.cc/paper/9364-debiased-bayesian-inference-for-average-treatment-effects>
AUTHORS: Kolyan Ray, Botond Szabo
HIGHLIGHT: Working in the standard potential outcomes framework, we propose a data-driven modification to an arbitrary (nonparametric) prior based on the propensity score that corrects for the first-order posterior bias, thereby improving performance.
- 1070, TITLE: Margin-Based Generalization Lower Bounds for Boosted Classifiers
<https://papers.nips.cc/paper/9365-margin-based-generalization-lower-bounds-for-boosted-classifiers>
AUTHORS: Allan Gr?nlund, Lior Kamma, Kasper Green Larsen, Alexander Mathiasen, Jelani Nelson
HIGHLIGHT: In this paper, we give the first margin-based lower bounds on the generalization error of boosted classifiers.
- 1071, TITLE: Connections Between Mirror Descent, Thompson Sampling and the Information Ratio
<https://papers.nips.cc/paper/9366-connections-between-mirror-descent-thompson-sampling-and-the-information-ratio>
AUTHORS: Julian Zimmert, Tor Lattimore
HIGHLIGHT: We make a formal connection, showing that the information-theoretic bounds in most applications are derived from existing techniques from online convex optimisation.
- 1072, TITLE: Graph Transformer Networks
<https://papers.nips.cc/paper/9367-graph-transformer-networks>
AUTHORS: Seongjun Yun, Minbyul Jeong, Raehyun Kim, Jaewoo Kang, Hyunwoo J. Kim
HIGHLIGHT: In this paper, we propose Graph Transformer Networks (GTNs) that are capable of generating new graph structures, which involve identifying useful connections between unconnected nodes on the original graph, while learning effective node representation on the new graphs in an end-to-end fashion.
- 1073, TITLE: Learning to Confuse: Generating Training Time Adversarial Data with Auto-Encoder
<https://papers.nips.cc/paper/9368-learning-to-confuse-generating-training-time-adversarial-data-with-auto-encoder>
AUTHORS: Ji Feng, Qi-Zhi Cai, Zhi-Hua Zhou
HIGHLIGHT: In this work, we consider one challenging training time attack by modifying training data with bounded perturbation, hoping to manipulate the behavior (both targeted or non-targeted) of any corresponding trained classifier during test time when facing clean samples.
- 1074, TITLE: The Impact of Regularization on High-dimensional Logistic Regression
<https://papers.nips.cc/paper/9369-the-impact-of-regularization-on-high-dimensional-logistic-regression>
AUTHORS: Fariborz Salehi, Ehsan Abbasi, Babak Hassibi
HIGHLIGHT: In the high-dimensional regime the underlying parameter vector is often structured (sparse, block-sparse, finite-alphabet, etc.) and so in this paper we study regularized logistic regression (RLR), where a convex regularizer that encourages the desired structure is added to the negative of the log-likelihood function.
- 1075, TITLE: Adaptive Density Estimation for Generative Models
<https://papers.nips.cc/paper/9370-adaptive-density-estimation-for-generative-models>
AUTHORS: Thomas Lucas, Konstantin Shmelkov, Karteek Alahari, Cordelia Schmid, Jakob Verbeek
HIGHLIGHT: As a solution, we propose the use of deep invertible transformations in the latent variable decoder.
- 1076, TITLE: Fast and Provable ADMM for Learning with Generative Priors
<https://papers.nips.cc/paper/9371-fast-and-provable-admm-for-learning-with-generative-priors>
AUTHORS: Fabian Latorre, Armin eftekhari, Volkan Cevher
HIGHLIGHT: In this work, we propose a (linearized) Alternating Direction Method-of-Multipliers (ADMM) algorithm for minimizing a convex function subject to a nonconvex constraint.
- 1077, TITLE: Weighted Linear Bandits for Non-Stationary Environments
<https://papers.nips.cc/paper/9372-weighted-linear-bandits-for-non-stationary-environments>
AUTHORS: Yoan Russac, Claire Vernade, Olivier Capp?
HIGHLIGHT: To address this problem, we propose D-LinUCB, a novel optimistic algorithm based on discounted linear regression, where exponential weights are used to smoothly forget the past.
- 1078, TITLE: Improved Regret Bounds for Bandit Combinatorial Optimization
<https://papers.nips.cc/paper/9373-improved-regret-bounds-for-bandit-combinatorial-optimization>
AUTHORS: Shinji Ito, Daisuke Hatano, Hanna Sumita, Kei Takemura, Takuro Fukunaga, Naonori Kakimura, Ken-Ichi Kawarabayashi

- HIGHLIGHT: In this paper, we aim to reveal the property, which makes the bandit combinatorial optimization hard.
- 1079, TITLE: Pareto Multi-Task Learning
<https://papers.nips.cc/paper/9374-pareto-multi-task-learning>
AUTHORS: Xi Lin, Hui-Ling Zhen, Zhenhua Li, Qing-Fu Zhang, Sam Kwong
HIGHLIGHT: In this paper, we generalize this idea and propose a novel Pareto multi-task learning algorithm (Pareto MTL) to find a set of well-distributed Pareto solutions which can represent different trade-offs among different tasks.
- 1080, TITLE: SIC-MMAB: Synchronisation Involves Communication in Multiplayer Multi-Armed Bandits
<https://papers.nips.cc/paper/9375-sic-mmab-synchronisation-involves-communication-in-multiplayer-multi-armed-bandits>
AUTHORS: Etienne Boursier, Vianney Perchet
HIGHLIGHT: We present a decentralized algorithm that achieves the same performance as a centralized one, contradicting the existing lower bounds for that problem.
- 1081, TITLE: Novel positional encodings to enable tree-based transformers
<https://papers.nips.cc/paper/9376-novel-positional-encodings-to-enable-tree-based-transformers>
AUTHORS: Vighnesh Shiv, Chris Quirk
HIGHLIGHT: Motivated by this property, we propose a method to extend transformers to tree-structured data, enabling sequence-to-tree, tree-to-sequence, and tree-to-tree mappings.
- 1082, TITLE: A Domain Agnostic Measure for Monitoring and Evaluating GANs
<https://papers.nips.cc/paper/9377-a-domain-agnostic-measure-for-monitoring-and-evaluating-gans>
AUTHORS: Paulina Grnarova, Kfir Y. Levy, Aurelien Lucchi, Nathanael Perraudin, Ian Goodfellow, Thomas Hofmann, Andreas Krause
HIGHLIGHT: We leverage the notion of duality gap from game theory to propose a measure that addresses both (i) and (ii) at a low computational cost.
- 1083, TITLE: Submodular Function Minimization with Noisy Evaluation Oracle
<https://papers.nips.cc/paper/9378-submodular-function-minimization-with-noisy-evaluation-oracle>
AUTHORS: Shinji Ito
HIGHLIGHT: For this problem, we provide an algorithm that returns an $O(n^{3/2}/\sqrt{T})$ -additive approximate solution in expectation, where n and T stand for the size of the problem and the number of oracle calls, respectively.
- 1084, TITLE: Counting the Optimal Solutions in Graphical Models
<https://papers.nips.cc/paper/9379-counting-the-optimal-solutions-in-graphical-models>
AUTHORS: Radu Marinescu, Rina Dechter
HIGHLIGHT: We introduce $\#opt$, a new inference task for graphical models which calls for counting the number of optimal solutions of the model.
- 1085, TITLE: Modelling the Dynamics of Multiagent Q-Learning in Repeated Symmetric Games: a Mean Field Theoretic Approach
<https://papers.nips.cc/paper/9380-modelling-the-dynamics-of-multiagent-q-learning-in-repeated-symmetric-games-a-mean-field-theoretic-approach>
AUTHORS: Shuyue Hu, Chin-wing Leung, Ho-fung Leung
HIGHLIGHT: In this paper, we study an n -agent setting with n tends to infinity, such that agents learn their policies concurrently over repeated symmetric bimatrix games with some other agents.
- 1086, TITLE: Deep Multimodal Multilinear Fusion with High-order Polynomial Pooling
<https://papers.nips.cc/paper/9381-deep-multimodal-multilinear-fusion-with-high-order-polynomial-pooling>
AUTHORS: Ming Hou, Jiajia Tang, Jianhai Zhang, Wanzeng Kong, Qibin Zhao
HIGHLIGHT: In this work, we first propose a polynomial tensor pooling (PTP) block for integrating multimodal features by considering high-order moments, followed by a tensorized fully connected layer. Treating PTP as a building block, we further establish a hierarchical polynomial fusion network (HPFN) to recursively transmit local correlations into global ones.
- 1087, TITLE: Bootstrapping Upper Confidence Bound
<https://papers.nips.cc/paper/9382-bootstrapping-upper-confidence-bound>
AUTHORS: Botao Hao, Yasin Abbasi Yadkori, Zheng Wen, Guang Cheng
HIGHLIGHT: In this paper, we propose a non-parametric and data-dependent UCB algorithm based on the multiplier bootstrap.

- 1088, TITLE: Integer Discrete Flows and Lossless Compression
<https://papers.nips.cc/paper/9383-integer-discrete-flows-and-lossless-compression>
AUTHORS: Emiel Hooeboom, Jorn Peters, Rianne van den Berg, Max Welling
HIGHLIGHT: For that reason, we introduce a flow-based generative model for ordinal discrete data called Integer Discrete Flow (IDF): a bijective integer map that can learn rich transformations on high-dimensional data.
- 1089, TITLE: Structured Prediction with Projection Oracles
<https://papers.nips.cc/paper/9384-structured-prediction-with-projection-oracles>
AUTHORS: Mathieu Blondel
HIGHLIGHT: We propose in this paper a general framework for deriving loss functions for structured prediction.
- 1090, TITLE: A Primal Dual Formulation For Deep Learning With Constraints
<https://papers.nips.cc/paper/9385-a-primal-dual-formulation-for-deep-learning-with-constraints>
AUTHORS: Yatin Nandwani, Abhishek Pathak, Mausam, Parag Singla
HIGHLIGHT: In this paper, we present a constrained optimization formulation for training a deep network with a given set of hard constraints on output labels.
- 1091, TITLE: Screening Sinkhorn Algorithm for Regularized Optimal Transport
<https://papers.nips.cc/paper/9386-screening-sinkhorn-algorithm-for-regularized-optimal-transport>
AUTHORS: Mokhtar Z. Alaya, Maxime Berar, Gilles Gasso, Alain Rakotomamonjy
HIGHLIGHT: We introduce in this paper a novel strategy for efficiently approximating the Sinkhorn distance between two discrete measures.
- 1092, TITLE: PAC-Bayes Un-Expected Bernstein Inequality
<https://papers.nips.cc/paper/9387-pac-bayes-un-expected-bernstein-inequality>
AUTHORS: Zakaria Mhammedi, Peter Gr?nwald, Benjamin Guedj
HIGHLIGHT: We present a new PAC-Bayesian generalization bound.
- 1093, TITLE: Are Labels Required for Improving Adversarial Robustness?
<https://papers.nips.cc/paper/9388-are-labels-required-for-improving-adversarial-robustness>
AUTHORS: Jean-Baptiste Alayrac, Jonathan Uesato, Po-Sen Huang, Alhussein Fawzi, Robert Stanforth, Pushmeet Kohli
HIGHLIGHT: Our main insight is that unlabeled data can be a competitive alternative to labeled data for training adversarially robust models.
- 1094, TITLE: Tight Regret Bounds for Model-Based Reinforcement Learning with Greedy Policies
<https://papers.nips.cc/paper/9389-tight-regret-bounds-for-model-based-reinforcement-learning-with-greedy-policies>
AUTHORS: Yonathan Efroni, Nadav Merlis, Mohammad Ghavamzadeh, Shie Mannor
HIGHLIGHT: In this paper, we focus on model-based RL in the finite-state finite-horizon MDP setting and establish that exploring with greedy policies -- act by 1-step planning -- can achieve tight minimax performance in terms of regret, $O(\sqrt{\text{HSAT}})$.
- 1095, TITLE: Multi-objective Bayesian optimisation with preferences over objectives
<https://papers.nips.cc/paper/9390-multi-objective-bayesian-optimisation-with-preferences-over-objectives>
AUTHORS: Majid Abdolshah, Alistair Shilton, Santu Rana, Sunil Gupta, Svetha Venkatesh
HIGHLIGHT: We present a multi-objective Bayesian optimisation algorithm that allows the user to express preference-order constraints on the objectives of the type objective A is more important than objective B.
- 1096, TITLE: Think out of the "Box": Generically-Constrained Asynchronous Composite Optimization and Hedging
<https://papers.nips.cc/paper/9391-think-out-of-the-box-generically-constrained-asynchronous-composite-optimization-and-hedging>
AUTHORS: Pooria Joulani, Andr?s Gy?rgy, Csaba Szepesvari
HIGHLIGHT: We present two new algorithms, ASYNCADA and HEDGEHOG, for asynchronous sparse online and stochastic optimization.
- 1097, TITLE: Calibration tests in multi-class classification: A unifying framework
<https://papers.nips.cc/paper/9392-calibration-tests-in-multi-class-classification-a-unifying-framework>
AUTHORS: David Widmann, Fredrik Lindsten, Dave Zachariah
HIGHLIGHT: We propose and study calibration measures for multi-class classification that generalize existing measures such as the expected calibration error, the maximum calibration error, and the maximum mean calibration error.
- 1098, TITLE: Classification Accuracy Score for Conditional Generative Models

<https://papers.nips.cc/paper/9393-classification-accuracy-score-for-conditional-generative-models>

AUTHORS: Suman Ravuri, Oriol Vinyals

HIGHLIGHT: To test this latter hypothesis, we use class-conditional generative models from a number of model classes - variational autoencoders, autoregressive models, and generative adversarial networks (GANs) - to infer the class labels of real data.

1099, TITLE: Theoretical Analysis of Adversarial Learning: A Minimax Approach

<https://papers.nips.cc/paper/9394-theoretical-analysis-of-adversarial-learning-a-minimax-approach>

AUTHORS: Zhuozhuo Tu, Jingwei Zhang, Dacheng Tao

HIGHLIGHT: In this paper, we propose a general theoretical method for analyzing the risk bound in the presence of adversaries.

1100, TITLE: Multiagent Evaluation under Incomplete Information

<https://papers.nips.cc/paper/9395-multiagent-evaluation-under-incomplete-information>

AUTHORS: Mark Rowland, Shayegan Omidshafiei, Karl Tuyls, Julien Perolat, Michal Valko, Georgios Piliouras, Remi Munos

HIGHLIGHT: We propose adaptive algorithms for accurate ranking, provide correctness and sample complexity guarantees, then introduce a means of connecting uncertainties in noisy match outcomes to uncertainties in rankings.

1101, TITLE: Tree-Sliced Variants of Wasserstein Distances

<https://papers.nips.cc/paper/9396-tree-sliced-variants-of-wasserstein-distances>

AUTHORS: Tam Le, Makoto Yamada, Kenji Fukumizu, Marco Cuturi

HIGHLIGHT: We consider in this work a more general family of ground metrics, namely ℓ_1 tree metrics, which also yield fast closed-form computations and negative definite, and of which the sliced-Wasserstein distance is a particular case (the tree is a chain).

1102, TITLE: Beyond temperature scaling: Obtaining well-calibrated multi-class probabilities with Dirichlet calibration

<https://papers.nips.cc/paper/9397-beyond-temperature-scaling-obtaining-well-calibrated-multi-class-probabilities-with-dirichlet-calibration>

AUTHORS: Meelis Kull, Miquel Perello Nieto, Markus K\u00f6ngsepp, Telmo Silva Filho, Hao Song, Peter Flach

HIGHLIGHT: We propose a natively multiclass calibration method applicable to classifiers from any model class, derived from Dirichlet distributions and generalising the beta calibration method from binary classification.

1103, TITLE: Comparing distributions: ℓ_1 geometry improves kernel two-sample testing

https://papers.nips.cc/paper/9398-comparing-distributions-ell_1-geometry-improves-kernel-two-sample-testing

AUTHORS: meyer scetbon, Gael Varoquaux

HIGHLIGHT: Here, we show that L^p distances (with $p \geq 1$) between these distribution representatives give metrics on the space of distributions that are well-behaved to detect differences between distributions as they metrize the weak convergence.

1104, TITLE: Robustness Verification of Tree-based Models

<https://papers.nips.cc/paper/9399-robustness-verification-of-tree-based-models>

AUTHORS: Hongge Chen, Huan Zhang, Si Si, Yang Li, Duane Boning, Cho-Jui Hsieh

HIGHLIGHT: For general problems, by exploiting the boxicity of the graph, we devise an efficient verification algorithm that can give tight lower bounds on robustness of decision tree ensembles, and allows iterative improvement and any-time termination.

1105, TITLE: Towards Interpretable Reinforcement Learning Using Attention Augmented Agents

<https://papers.nips.cc/paper/9400-towards-interpretable-reinforcement-learning-using-attention-augmented-agents>

AUTHORS: Alexander Mott, Daniel Zoran, Mike Chrzanowski, Daan Wierstra, Danilo Jimenez Rezende

HIGHLIGHT: Inspired by recent work in attention models for image captioning and question answering, we present a soft attention model for the reinforcement learning domain.

1106, TITLE: Fast and Accurate Stochastic Gradient Estimation

<https://papers.nips.cc/paper/9401-fast-and-accurate-stochastic-gradient-estimation>

AUTHORS: Beidi Chen, Yingchen Xu, Anshumali Shrivastava

HIGHLIGHT: In this paper, we break this barrier by providing the first demonstration of a scheme, Locality sensitive hashing (LSH) sampled Stochastic Gradient Descent (LGD), which leads to superior gradient estimation while keeping the sampling cost per iteration similar to that of the uniform sampling.

1107, TITLE: Theoretical Limits of Pipeline Parallel Optimization and Application to Distributed Deep Learning

<https://papers.nips.cc/paper/9402-theoretical-limits-of-pipeline-parallel-optimization-and-application-to-distributed-deep-learning>

AUTHORS: Igor Colin, Ludovic DOS SANTOS, Kevin Scaman

HIGHLIGHT: We investigate the theoretical limits of pipeline parallel learning of deep learning architectures, a distributed setup in which the computation is distributed per layer instead of per example.

1108, **TITLE:** Root Mean Square Layer Normalization
<https://papers.nips.cc/paper/9403-root-mean-square-layer-normalization>

AUTHORS: Biao Zhang, Rico Sennrich

HIGHLIGHT: In this paper, we hypothesize that re-centering invariance in LayerNorm is dispensable and propose root mean square layer normalization, or RMSNorm.

1109, **TITLE:** Universality in Learning from Linear Measurements
<https://papers.nips.cc/paper/9404-universality-in-learning-from-linear-measurements>

AUTHORS: Ehsan Abbasi, Fariborz Salehi, Babak Hassibi

HIGHLIGHT: We study the problem of recovering a structured signal from independently and identically drawn linear measurements.

1110, **TITLE:** Planning in entropy-regularized Markov decision processes and games
<https://papers.nips.cc/paper/9405-planning-in-entropy-regularized-markov-decision-processes-and-games>

AUTHORS: Jean-Bastien Grill, Omar Darwiche Domingues, Pierre Menard, Remi Munos, Michal Valko

HIGHLIGHT: We propose SmoothCruiser, a new planning algorithm for estimating the value function in entropy-regularized Markov decision processes and two-player games, given a generative model of the SmoothCruiser.

1111, **TITLE:** Exponentially convergent stochastic k-PCA without variance reduction

<https://papers.nips.cc/paper/9406-exponentially-convergent-stochastic-k-pca-without-variance-reduction>

AUTHORS: Cheng Tang

HIGHLIGHT: We present Matrix Krasulina, an algorithm for online k-PCA, by generalizing the classic Krasulina's method (Krasulina, 1969) from vector to matrix case.

1112, **TITLE:** R2D2: Reliable and Repeatable Detector and Descriptor

<https://papers.nips.cc/paper/9407-r2d2-reliable-and-repeatable-detector-and-descriptor>

AUTHORS: Jerome Revaud, Cesar De Souza, Martin Humenberger, Philippe Weinzaepfel

HIGHLIGHT: In this work, we argue that repeatable regions are not necessarily discriminative and can therefore lead to select suboptimal keypoints.

1113, **TITLE:** Selective Sampling-based Scalable Sparse Subspace Clustering

<https://papers.nips.cc/paper/9408-selective-sampling-based-scalable-sparse-subspace-clustering>

AUTHORS: Shin Matsushima, Maria Brbic

HIGHLIGHT: To overcome this limitation, we introduce Selective Sampling-based Scalable Sparse Subspace Clustering (SSC) algorithm which selects subsamples based on the approximated subgradients and linearly scales with the number of data points in terms of time and memory requirements.

1114, **TITLE:** A General Framework for Symmetric Property Estimation

<https://papers.nips.cc/paper/9409-a-general-framework-for-symmetric-property-estimation>

AUTHORS: Moses Charikar, Kirankumar Shiragur, Aaron Sidford

HIGHLIGHT: In this paper we provide a general framework for estimating symmetric properties of distributions from i.i.d. samples.

1115, **TITLE:** Structured Variational Inference in Continuous Cox Process Models

<https://papers.nips.cc/paper/9410-structured-variational-inference-in-continuous-cox-process-models>

AUTHORS: Virginia Aglietti, Edwin V. Bonilla, Theodoros Damoulas, Sally Cripps

HIGHLIGHT: We propose a scalable framework for inference in a continuous sigmoidal Cox process that assumes the corresponding intensity function is given by a Gaussian process (GP) prior transformed with a scaled logistic sigmoid function.

1116, **TITLE:** Generalization of Reinforcement Learners with Working and Episodic Memory

<https://papers.nips.cc/paper/9411-generalization-of-reinforcement-learners-with-working-and-episodic-memory>

AUTHORS: Meire Fortunato, Melissa Tan, Ryan Faulkner, Steven Hansen, Adrià Puigdomènech Badia, Gavin Buttmore, Charles Deck, Joel Z. Leibo, Charles Blundell

HIGHLIGHT: In this paper, we aim to develop a comprehensive methodology to test different kinds of memory in an agent and assess how well the agent can apply what it learns in training to a holdout set that differs from the training set along dimensions that we suggest are relevant for evaluating memory-specific generalization.

1117, TITLE: Distribution Learning of a Random Spatial Field with a Location-Unaware Mobile Sensor
<https://papers.nips.cc/paper/9412-distribution-learning-of-a-random-spatial-field-with-a-location-unaware-mobile-sensor>
AUTHORS: Meera Pai, Animesh Kumar
HIGHLIGHT: While GPS or other localization methods can reduce this uncertainty, we address a more fundamental question: can a location-unaware mobile sensor, recording samples on a directed non-uniform random walk, learn the statistical distribution (as a function of space) of an underlying random process (spatial field)?

1118, TITLE: Hindsight Credit Assignment
<https://papers.nips.cc/paper/9413-hindsight-credit-assignment>
AUTHORS: Anna Harutyunyan, Will Dabney, Thomas Mesnard, Mohammad Gheshlaghi Azar, Bilal Piot, Nicolas Heess, Hado P. van Hasselt, Gregory Wayne, Satinder Singh, Doina Precup, Remi Munos
HIGHLIGHT: We consider the problem of efficient credit assignment in reinforcement learning.

1119, TITLE: Efficient Identification in Linear Structural Causal Models with Instrumental Cutsets
<https://papers.nips.cc/paper/9414-efficient-identification-in-linear-structural-causal-models-with-instrumental-cutsets>
AUTHORS: Daniel Kumor, Bryant Chen, Elias Bareinboim
HIGHLIGHT: In this paper, we investigate graphical conditions to allow efficient identification in arbitrary linear structural causal models (SCMs).

1120, TITLE: Kernelized Bayesian Softmax for Text Generation
<https://papers.nips.cc/paper/9415-kernelized-bayesian-softmax-for-text-generation>
AUTHORS: Ning Miao, Hao Zhou, Chengqi Zhao, Wenxian Shi, Lei Li
HIGHLIGHT: In this paper, we propose KerBS, a novel approach for learning better embeddings for text generation.

1121, TITLE: When to Trust Your Model: Model-Based Policy Optimization
<https://papers.nips.cc/paper/9416-when-to-trust-your-model-model-based-policy-optimization>
AUTHORS: Michael Janner, Justin Fu, Marvin Zhang, Sergey Levine
HIGHLIGHT: In this paper, we study the role of model usage in policy optimization both theoretically and empirically.

1122, TITLE: Correlation Clustering with Adaptive Similarity Queries
<https://papers.nips.cc/paper/9417-correlation-clustering-with-adaptive-similarity-queries>
AUTHORS: Marco Bressan, Nicolò Cesa-Bianchi, Andrea Paudice, Fabio Vitale
HIGHLIGHT: In this work we investigate correlation clustering as an active learning problem: each similarity score can be learned by making a query, and the goal is to minimise both the disagreements and the total number of queries.

1123, TITLE: Control What You Can: Intrinsically Motivated Task-Planning Agent
<https://papers.nips.cc/paper/9418-control-what-you-can-intrinsically-motivated-task-planning-agent>
AUTHORS: Sebastian Blaes, Marin Vlastelica Pogancic, Jiajie Zhu, Georg Martius
HIGHLIGHT: We present a novel intrinsically motivated agent that learns how to control the environment in a sample efficient manner, that is with as few environment interactions as possible, by optimizing learning progress.

1124, TITLE: Selecting causal brain features with a single conditional independence test per feature
<https://papers.nips.cc/paper/9419-selecting-causal-brain-features-with-a-single-conditional-independence-test-per-feature>
AUTHORS: Atalanti Mastakouri, Bernhard Schölkopf, Dominik Janzing
HIGHLIGHT: We propose a constraint-based causal feature selection method for identifying causes of a given target variable, selecting from a set of candidate variables, while there can also be hidden variables acting as common causes with the target.

1125, TITLE: Continuous Hierarchical Representations with Poincaré Variational Auto-Encoders
<https://papers.nips.cc/paper/9420-continuous-hierarchical-representations-with-poincare-variational-auto-encoders>
AUTHORS: Emile Mathieu, Charline Le Lan, Chris J. Maddison, Ryota Tomioka, Yee Whye Teh
HIGHLIGHT: We therefore endow VAEs with a Poincaré ball model of hyperbolic geometry as a latent space and rigorously derive the necessary methods to work with two main Gaussian generalisations on that space.

1126, TITLE: A Generic Acceleration Framework for Stochastic Composite Optimization
<https://papers.nips.cc/paper/9421-a-generic-acceleration-framework-for-stochastic-composite-optimization>
AUTHORS: Andrei Kulunchakov, Julien Mairal
HIGHLIGHT: In this paper, we introduce various mechanisms to obtain accelerated first-order stochastic optimization algorithms when the objective function is convex or strongly convex.

1127, TITLE: Beating SGD Saturation with Tail-Averaging and Minibatching

<https://papers.nips.cc/paper/9422-beating-sgd-saturation-with-tail-averaging-and-minibatching>
AUTHORS: Nicole Muecke, Gergely Neu, Lorenzo Rosasco
HIGHLIGHT: In this paper, we consider least squares learning in a nonparametric setting and contribute to filling this gap by focusing on the effect and interplay of multiple passes, mini-batching and averaging, in particular tail averaging.

1128, TITLE: Random Quadratic Forms with Dependence: Applications to Restricted Isometry and Beyond
<https://papers.nips.cc/paper/9423-random-quadratic-forms-with-dependence-applications-to-restricted-isometry-and-beyond>
AUTHORS: Arindam Banerjee, Qilong Gu, Vidyashankar Sivakumar, Steven Z. Wu
HIGHLIGHT: In this paper, we show that such independence is in fact not needed for such results which continue to hold under fairly general dependence structures.

1129, TITLE: Continuous-time Models for Stochastic Optimization Algorithms
<https://papers.nips.cc/paper/9424-continuous-time-models-for-stochastic-optimization-algorithms>
AUTHORS: Antonio Orvieto, Aurelien Lucchi
HIGHLIGHT: We propose new continuous-time formulations for first-order stochastic optimization algorithms such as mini-batch gradient descent and variance-reduced methods.

1130, TITLE: Curriculum-guided Hindsight Experience Replay
<https://papers.nips.cc/paper/9425-curriculum-guided-hindsight-experience-replay>
AUTHORS: Meng Fang, Tianyi Zhou, Yali Du, Lei Han, Zhengyou Zhang
HIGHLIGHT: In this paper, we propose to 1) adaptively select the failed experiences for replay according to the proximity to the true goals and the curiosity of exploration over diverse pseudo goals, and 2) gradually change the proportion of the goal-proximity and the diversity-based curiosity in the selection criteria: we adopt a human-like learning strategy that enforces more curiosity in earlier stages and changes to larger goal-proximity later.

1131, TITLE: Implicit Semantic Data Augmentation for Deep Networks
<https://papers.nips.cc/paper/9426-implicit-semantic-data-augmentation-for-deep-networks>
AUTHORS: Yulin Wang, Xuran Pan, Shiji Song, Hong Zhang, Gao Huang, Cheng Wu
HIGHLIGHT: In this paper, we propose a novel implicit semantic data augmentation (ISDA) approach to complement traditional augmentation techniques like flipping, translation or rotation.

1132, TITLE: MetaInit: Initializing learning by learning to initialize
<https://papers.nips.cc/paper/9427-metainit-initializing-learning-by-learning-to-initialize>
AUTHORS: Yann N. Dauphin, Samuel Schoenholz
HIGHLIGHT: In this work, we introduce an algorithm called MetaInit as a step towards automating the search for good initializations using meta-learning.

1133, TITLE: Scalable Deep Generative Relational Model with High-Order Node Dependence
<https://papers.nips.cc/paper/9428-scalable-deep-generative-relational-model-with-high-order-node-dependence>
AUTHORS: Xuhui Fan, Bin Li, Caoyuan Li, Scott Sisson, Ling Chen
HIGHLIGHT: In this work, we propose a probabilistic framework for relational data modelling and latent structure exploring.

1134, TITLE: Random Path Selection for Continual Learning
<https://papers.nips.cc/paper/9429-random-path-selection-for-continual-learning>
AUTHORS: Jathushan Rajasegaran, Munawar Hayat, Salman H. Khan, Fahad Shahbaz Khan, Ling Shao
HIGHLIGHT: In this paper, we propose a random path selection algorithm, called RPS-Net, that progressively chooses optimal paths for the new tasks while encouraging parameter sharing and reuse.

1135, TITLE: Efficient Algorithms for Smooth Minimax Optimization
<https://papers.nips.cc/paper/9430-efficient-algorithms-for-smooth-minimax-optimization>
AUTHORS: Kiran K. Thekumparampil, Prateek Jain, Praneeth Netrapalli, Sewoong Oh
HIGHLIGHT: For strongly-convex $g(\cdot, y), \forall y$, we propose a new direct optimal algorithm combining Mirror-Prox and Nesterov's AGD, and show that it can find global optimum in $\widetilde{O}\left(\frac{1}{k^2}\right)$ iterations, improving over current state-of-the-art rate of $O\left(\frac{1}{k}\right)$.

1136, TITLE: Shadowing Properties of Optimization Algorithms
<https://papers.nips.cc/paper/9431-shadowing-properties-of-optimization-algorithms>
AUTHORS: Antonio Orvieto, Aurelien Lucchi
HIGHLIGHT: In an attempt to encourage the use of continuous-time methods in optimization, we show that, if some additional regularity on the objective is assumed, the ODE representations of Gradient Descent and Heavy-ball do not suffer from the aforementioned problem, once we allow for a small perturbation on the algorithm initial condition.

1137, TITLE: Causal Regularization
<https://papers.nips.cc/paper/9432-causal-regularization>
AUTHORS: Dominik Janzing
HIGHLIGHT: We argue that regularizing terms in standard regression methods not only help against overfitting finite data, but sometimes also help in getting better causal models.

1138, TITLE: Learning Hawkes Processes from a handful of events
<https://papers.nips.cc/paper/9433-learning-hawkes-processes-from-a-handful-of-events>
AUTHORS: Farnood Salehi, William Trouleau, Matthias Grossglauser, Patrick Thiran
HIGHLIGHT: To solve both issues, we develop in this work an efficient algorithm based on variational expectation-maximization.

1139, TITLE: Unsupervised Object Segmentation by Redrawing
<https://papers.nips.cc/paper/9434-unsupervised-object-segmentation-by-redrawing>
AUTHORS: Micka?l Chen, Thierry Arti?res, Ludovic Denoyer
HIGHLIGHT: We present ReDO, a new model able to extract objects from images without any annotation in an unsupervised way.

1140, TITLE: Regret Bounds for Learning State Representations in Reinforcement Learning
<https://papers.nips.cc/paper/9435-regret-bounds-for-learning-state-representations-in-reinforcement-learning>
AUTHORS: Ronald Ortner, Matteo Pirota, Alessandro Lazaric, Ronan Fruit, Odalric-Ambrym Maillard
HIGHLIGHT: We propose an algorithm (UCB-MS) with $O(\sqrt{T})$ regret in any communicating Markov decision process.

1141, TITLE: Band-Limited Gaussian Processes: The Sinc Kernel
<https://papers.nips.cc/paper/9436-band-limited-gaussian-processes-the-sinc-kernel>
AUTHORS: Felipe Tobar
HIGHLIGHT: We propose a novel class of Gaussian processes (GPs) whose spectra have compact support, meaning that their sample trajectories are almost-surely band limited.

1142, TITLE: Leveraging Labeled and Unlabeled Data for Consistent Fair Binary Classification
<https://papers.nips.cc/paper/9437-leveraging-labeled-and-unlabeled-data-for-consistent-fair-binary-classification>
AUTHORS: Evgenii Chzhen, Christophe Denis, Mohamed Hebiri, Luca Oneto, Massimiliano Pontil
HIGHLIGHT: We study the problem of fair binary classification using the notion of Equal Opportunity.

1143, TITLE: Learning search spaces for Bayesian optimization: Another view of hyperparameter transfer learning
<https://papers.nips.cc/paper/9438-learning-search-spaces-for-bayesian-optimization-another-view-of-hyperparameter-transfer-learning>
AUTHORS: Valerio Perrone, Huibin Shen
HIGHLIGHT: In this work, we introduce a method to automatically design the BO search space by relying on evaluations of previous black-box functions.

1144, TITLE: Streaming Bayesian Inference for Crowdsourced Classification
<https://papers.nips.cc/paper/9439-streaming-bayesian-inference-for-crowdsourced-classification>
AUTHORS: Edoardo Manino, Long Tran-Thanh, Nicholas Jennings
HIGHLIGHT: In this paper, we revisit the problem of binary classification from crowdsourced data.

1145, TITLE: Neuropathic Pain Diagnosis Simulator for Causal Discovery Algorithm Evaluation
<https://papers.nips.cc/paper/9440-neuropathic-pain-diagnosis-simulator-for-causal-discovery-algorithm-evaluation>
AUTHORS: Ruibo Tu, Kun Zhang, Bo Bertilson, Hedvig Kjellstrom, Cheng Zhang
HIGHLIGHT: In this work, we handle the problem of evaluating causal discovery algorithms by building a flexible simulator in the medical setting.
We develop a neuropathic pain diagnosis simulator, inspired by the fact that the biological processes of neuropathic pathophysiology are well studied with well-understood causal influences.

1146, TITLE: Brain-Like Object Recognition with High-Performing Shallow Recurrent ANNs
<https://papers.nips.cc/paper/9441-brain-like-object-recognition-with-high-performing-shallow-recurrent-anns>
AUTHORS: Jonas Kubilius, Martin Schrimpf, Ha Hong, Najib Majaj, Rishi Rajalingham, Elias Issa, Kohitij Kar, Pouya Bashivan, Jonathan Prescott-Roy, Kailyn Schmidt, Aran Nayebi, Daniel Bear, Daniel L. Yamins, James J. DiCarlo
HIGHLIGHT: Here we demonstrate that better anatomical alignment to the brain and high performance on machine learning as well as neuroscience measures do not have to be in contradiction.

- 1147, TITLE: k-Means Clustering of Lines for Big Data
<https://papers.nips.cc/paper/9442-k-means-clustering-of-lines-for-big-data>
AUTHORS: Yair Marom, Dan Feldman
HIGHLIGHT: We suggest the first PTAS that computes a $(1+\epsilon)$ -approximation to this problem in time $O(n \log n)$ for any constant approximation error $\epsilon \in (0, 1)$, and constant integers $k, d \geq 1$.
- 1148, TITLE: Random Projections and Sampling Algorithms for Clustering of High-Dimensional Polygonal Curves
<https://papers.nips.cc/paper/9443-random-projections-and-sampling-algorithms-for-clustering-of-high-dimensional-polygonal-curves>
AUTHORS: Stefan Meintrup, Alexander Munteanu, Dennis Rohde
HIGHLIGHT: We study the k -median clustering problem for high-dimensional polygonal curves with finite but unbounded number of vertices.
- 1149, TITLE: Recurrent Space-time Graph Neural Networks
<https://papers.nips.cc/paper/9444-recurrent-space-time-graph-neural-networks>
AUTHORS: Andrei Nicolicioiu, Iulia Duta, Marius Leordeanu
HIGHLIGHT: We propose a neural graph model, recurrent in space and time, suitable for capturing both the local appearance and the complex higher-level interactions of different entities and objects within the changing world scene.
- 1150, TITLE: Uncertainty on Asynchronous Time Event Prediction
<https://papers.nips.cc/paper/9445-uncertainty-on-asynchronous-time-event-prediction>
AUTHORS: Bertrand Charpentier, Marin Bilou, Stephan G?nnemann
HIGHLIGHT: In this work, we tackle the task of predicting the next event (given a history), and how this prediction changes with the passage of time.
- 1151, TITLE: Accurate, reliable and fast robustness evaluation
<https://papers.nips.cc/paper/9446-accurate-reliable-and-fast-robustness-evaluation>
AUTHORS: Wieland Brendel, Jonas Rauber, Matthias K?mmerer, Ivan Ustuzhaninov, Matthias Bethge
HIGHLIGHT: We here develop a new set of gradient-based adversarial attacks which (a) are more reliable in the face of gradient-masking than other gradient-based attacks, (b) perform better and are more query efficient than current state-of-the-art gradient-based attacks, (c) can be flexibly adapted to a wide range of adversarial criteria and (d) require virtually no hyperparameter tuning.
- 1152, TITLE: Sparse High-Dimensional Isotonic Regression
<https://papers.nips.cc/paper/9447-sparse-high-dimensional-isotonic-regression>
AUTHORS: David Gamamik, Julia Gaudio
HIGHLIGHT: We consider the problem of estimating an unknown coordinate-wise monotone function given noisy measurements, known as the isotonic regression problem.
- 1153, TITLE: Triad Constraints for Learning Causal Structure of Latent Variables
<https://papers.nips.cc/paper/9448-triad-constraints-for-learning-causal-structure-of-latent-variables>
AUTHORS: Ruichu Cai, Feng Xie, Clark Glymour, Zhifeng Hao, Kun Zhang
HIGHLIGHT: In this paper, by properly leveraging the non-Gaussianity of the data, we propose to estimate the structure over latent variables with the so-called Triad constraints: we design a form of "pseudo-residual" from three variables, and show that when causal relations are linear and noise terms are non-Gaussian, the causal direction between the latent variables for the three observed variables is identifiable by checking a certain kind of independence relationship.
- 1154, TITLE: On the Inductive Bias of Neural Tangent Kernels
<https://papers.nips.cc/paper/9449-on-the-inductive-bias-of-neural-tangent-kernels>
AUTHORS: Alberto Bietti, Julien Mairal
HIGHLIGHT: In particular, we study smoothness, approximation, and stability properties of functions with finite norm, including stability to image deformations in the case of convolutional networks, and compare to other known kernels for similar architectures.
- 1155, TITLE: Cross-Domain Transferability of Adversarial Perturbations
<https://papers.nips.cc/paper/9450-cross-domain-transferability-of-adversarial-perturbations>
AUTHORS: Muhammad Muzammal Naseer, Salman H. Khan, Muhammad Haris Khan, Fahad Shahbaz Khan, Fatih Porikli
HIGHLIGHT: To this end, we propose a framework capable of launching highly transferable attacks that crafts adversarial patterns to mislead networks trained on wholly different domains.

- 1156, TITLE: Shallow RNN: Accurate Time-series Classification on Resource Constrained Devices
<https://papers.nips.cc/paper/9451-shallow-rnn-accurate-time-series-classification-on-resource-constrained-devices>
AUTHORS: Don Dennis, Durmus Alp Emre Acar, Vikram Mandikal, Vinu Sankar Sadasivan, Venkatesh Saligrama, Harsha Vardhan Simhadri, Prateek Jain
HIGHLIGHT: To induce long-term dependencies, and yet admit parallelization, we introduce novel shallow RNNs.
- 1157, TITLE: Kernel quadrature with DPPs
<https://papers.nips.cc/paper/9452-kernel-quadrature-with-dpps>
AUTHORS: Ayoub Belhadji, R?mi Bardenet, Pierre Chainais
HIGHLIGHT: We study quadrature rules for functions living in an RKHS, using nodes sampled from a projection determinantal point process (DPP).
- 1158, TITLE: REM: From Structural Entropy to Community Structure Deception
<https://papers.nips.cc/paper/9453-rem-from-structural-entropy-to-community-structure-deception>
AUTHORS: Yiwei Liu, Jiamou Liu, Zijian Zhang, Liehuang Zhu, Angsheng Li
HIGHLIGHT: To this end, we propose a community-based structural entropy to express the amount of information revealed by a community structure.
- 1159, TITLE: Sim2real transfer learning for 3D human pose estimation: motion to the rescue
<https://papers.nips.cc/paper/9454-sim2real-transfer-learning-for-3d-human-pose-estimation-motion-to-the-rescue>
AUTHORS: Carl Doersch, Andrew Zisserman
HIGHLIGHT: In this paper, we show that standard neural-network approaches, which perform poorly when trained on synthetic RGB images, can perform well when the data is pre-processed to extract cues about the person's motion, notably as optical flow and the motion of 2D keypoints.
- 1160, TITLE: Self-Supervised Deep Learning on Point Clouds by Reconstructing Space
<https://papers.nips.cc/paper/9455-self-supervised-deep-learning-on-point-clouds-by-reconstructing-space>
AUTHORS: Bjame Sievers, Jonathan Sauder
HIGHLIGHT: We propose a self-supervised learning task for deep learning on raw point cloud data in which a neural network is trained to reconstruct point clouds whose parts have been randomly rearranged.
- 1161, TITLE: Piecewise Strong Convexity of Neural Networks
<https://papers.nips.cc/paper/9456-piecewise-strong-convexity-of-neural-networks>
AUTHORS: Tristan Milne
HIGHLIGHT: We study the loss surface of a feed-forward neural network with ReLU non-linearities, regularized with weight decay.
- 1162, TITLE: Minimum Stein Discrepancy Estimators
<https://papers.nips.cc/paper/9457-minimum-stein-discrepancy-estimators>
AUTHORS: Alessandro Barp, Francois-Xavier Briol, Andrew Duncan, Mark Girolami, Lester Mackey
HIGHLIGHT: We provide a unifying perspective of these techniques as minimum Stein discrepancy estimators, and use this lens to design new diffusion kernel Stein discrepancy (DKSD) and diffusion score matching (DSM) estimators with complementary strengths.
- 1163, TITLE: Fast and Furious Learning in Zero-Sum Games: Vanishing Regret with Non-Vanishing Step Sizes
<https://papers.nips.cc/paper/9458-fast-and-furious-learning-in-zero-sum-games-vanishing-regret-with-non-vanishing-step-sizes>
AUTHORS: James Bailey, Georgios Piliouras
HIGHLIGHT: We show for the first time that it is possible to reconcile in online learning in zero-sum games two seemingly contradictory objectives: vanishing time-average regret and non-vanishing step sizes.
- 1164, TITLE: Generalization Bounds for Neural Networks via Approximate Description Length
<https://papers.nips.cc/paper/9459-generalization-bounds-for-neural-networks-via-approximate-description-length>
AUTHORS: Amit Daniely, Elad Granot
HIGHLIGHT: We investigate the sample complexity of networks with bounds on the magnitude of its weights.
- 1165, TITLE: Provably robust boosted decision stumps and trees against adversarial attacks
<https://papers.nips.cc/paper/9460-provably-robust-boosted-decision-stumps-and-trees-against-adversarial-attacks>
AUTHORS: Maksym Andriushchenko, Matthias Hein
HIGHLIGHT: We show in this paper that for boosted decision stumps the min-max robust loss and test error for an ℓ_∞ -attack can be computed in $\mathcal{O}(T \log T)$ time per input, where T is the number of decision stumps and the optimal update step of the ensemble can be done in $\mathcal{O}(n^2, T \log T)$, where n is the number of data points.

- 1166, TITLE: Convergence of Adversarial Training in Overparametrized Neural Networks
<https://papers.nips.cc/paper/9461-convergence-of-adversarial-training-in-overparametrized-neural-networks>
AUTHORS: Ruiqi Gao, Tianle Cai, Haochuan Li, Cho-Jui Hsieh, Liwei Wang, Jason D. Lee
HIGHLIGHT: This paper provides a partial answer to the success of adversarial training, by showing that it converges to a network where the surrogate loss with respect to the the attack algorithm is within ϵ of the optimal robust loss.
- 1167, TITLE: A Composable Specification Language for Reinforcement Learning Tasks
<https://papers.nips.cc/paper/9462-a-composable-specification-language-for-reinforcement-learning-tasks>
AUTHORS: Kishor Jothimurugan, Rajeev Alur, Osbert Bastani
HIGHLIGHT: We propose a language for specifying complex control tasks, along with an algorithm that compiles specifications in our language into a reward function and automatically performs reward shaping.
- 1168, TITLE: The Option Keyboard: Combining Skills in Reinforcement Learning
<https://papers.nips.cc/paper/9463-the-option-keyboard-combining-skills-in-reinforcement-learning>
AUTHORS: Andre Barreto, Diana Borsa, Shaobo Hou, Gheorghe Comanici, Eser Aygün, Philippe Hamel, Daniel Toyama, Jonathan Hunt, Shihab Mourad, David Silver, Doina Precup
HIGHLIGHT: Based on this premise, we propose a framework for combining skills using the formalism of options.
- 1169, TITLE: Unified Language Model Pre-training for Natural Language Understanding and Generation
<https://papers.nips.cc/paper/9464-unified-language-model-pre-training-for-natural-language-understanding-and-generation>
AUTHORS: Li Dong, Nan Yang, Wenhui Wang, Furu Wei, Xiaodong Liu, Yu Wang, Jianfeng Gao, Ming Zhou, Hsiao-Wen Hon
HIGHLIGHT: This paper presents a new Unified pre-trained Language Model (UniLM) that can be fine-tuned for both natural language understanding and generation tasks.
- 1170, TITLE: Learning to Correlate in Multi-Player General-Sum Sequential Games
<https://papers.nips.cc/paper/9465-learning-to-correlate-in-multi-player-general-sum-sequential-games>
AUTHORS: Andrea Celli, Alberto Marchesi, Tommaso Bianchi, Nicola Gatti
HIGHLIGHT: In this paper, we focus on coarse correlated equilibria (CCEs) in sequential games.
- 1171, TITLE: Stochastic Continuous Greedy ++: When Upper and Lower Bounds Match
<https://papers.nips.cc/paper/9466-stochastic-continuous-greedy-when-upper-and-lower-bounds-match>
AUTHORS: Amin Karbasi, Hamed Hassani, Aryan Mokhtari, Zebang Shen
HIGHLIGHT: In this paper, we develop $\text{scg}^+(\text{SCG}^{\{++\}})$, the first efficient variant of a conditional gradient method for maximizing a continuous submodular function subject to a convex constraint.
- 1172, TITLE: Generative Well-intentioned Networks
<https://papers.nips.cc/paper/9467-generative-well-intentioned-networks>
AUTHORS: Justin Cosentino, Jun Zhu
HIGHLIGHT: We propose Generative Well-intentioned Networks (GWINs), a novel framework for increasing the accuracy of certainty-based, closed-world classifiers.
- 1173, TITLE: Online-Within-Online Meta-Learning
<https://papers.nips.cc/paper/9468-online-within-online-meta-learning>
AUTHORS: Giulia Denevi, Dimitris Stamos, Carlo Ciliberto, Massimiliano Pontil
HIGHLIGHT: We study the problem of learning a series of tasks in a fully online Meta-Learning setting.
- 1174, TITLE: Learning step sizes for unfolded sparse coding
<https://papers.nips.cc/paper/9469-learning-step-sizes-for-unfolded-sparse-coding>
AUTHORS: Pierre Ablin, Thomas Moreau, Mathurin Massias, Alexandre Gramfort
HIGHLIGHT: In this paper, we study the selection of adapted step sizes for ISTA.
- 1175, TITLE: Biases for Emergent Communication in Multi-agent Reinforcement Learning
<https://papers.nips.cc/paper/9470-biases-for-emergent-communication-in-multi-agent-reinforcement-learning>
AUTHORS: Tom Eccles, Yoram Bachrach, Guy Lever, Angeliki Lazaridou, Thore Graepel
HIGHLIGHT: We introduce inductive biases for positive signalling and positive listening, which ease this problem.
- 1176, TITLE: Episodic Memory in Lifelong Language Learning

<https://papers.nips.cc/paper/9471-episodic-memory-in-lifelong-language-learning>
AUTHORS: Cyprien de Masson d'Autume, Sebastian Ruder, Lingpeng Kong, Dani Yogatama
HIGHLIGHT: We propose an episodic memory model that performs sparse experience replay and local adaptation to mitigate catastrophic forgetting in this setup.

1177, TITLE: A Simple Baseline for Bayesian Uncertainty in Deep Learning
<https://papers.nips.cc/paper/9472-a-simple-baseline-for-bayesian-uncertainty-in-deep-learning>
AUTHORS: Wesley J. Maddox, Pavel Izmailov, Timur Garipov, Dmitry P. Vetrov, Andrew Gordon Wilson
HIGHLIGHT: We propose SWA-Gaussian (SWAG), a simple, scalable, and general purpose approach for uncertainty representation and calibration in deep learning.

1178, TITLE: Communication-efficient Distributed SGD with Sketching
<https://papers.nips.cc/paper/9473-communication-efficient-distributed-sgd-with-sketching>
AUTHORS: Nikita Ivkin, Daniel Rothchild, Enayat Ullah, Vladimir Braverman, Ion Stoica, Raman Arora
HIGHLIGHT: Motivated by the success of sketching methods in sub-linear/streaming algorithms, we introduce Sketched-SGD, an algorithm for carrying out distributed SGD by communicating sketches instead of full gradients.

1179, TITLE: Modeling Conceptual Understanding in Image Reference Games
<https://papers.nips.cc/paper/9474-modeling-conceptual-understanding-in-image-reference-games>
AUTHORS: Rodolfo Corona Rodriguez, Stephan Alaniz, Zeynep Akata
HIGHLIGHT: In this work, we present both an image reference game between a speaker and a population of listeners where reasoning about the concepts other agents can comprehend is necessary and a model formulation with this capability.

1180, TITLE: Kalman Filter, Sensor Fusion, and Constrained Regression: Equivalences and Insights
<https://papers.nips.cc/paper/9475-kalman-filter-sensor-fusion-and-constrained-regression-equivalences-and-insights>
AUTHORS: Maria Jahja, David Farrow, Roni Rosenfeld, Ryan J. Tibshirani
HIGHLIGHT: In this work, we show that the state estimates from the KF in a standard linear dynamical system setting are equivalent to those given by the KF in a transformed system, with infinite process noise (i.e., a "flat prior") and an augmented measurement space.

1181, TITLE: Near Neighbor: Who is the Fairest of Them All?
<https://papers.nips.cc/paper/9476-near-neighbor-who-is-the-fairest-of-them-all>
AUTHORS: Sarel Har-Peled, Sepideh Mahabadi
HIGHLIGHT: In this work we study a "fair" variant of the near neighbor problem.

1182, TITLE: Outlier-robust estimation of a sparse linear model using ℓ_1 -penalized Huber-M-estimator
https://papers.nips.cc/paper/9477-outlier-robust-estimation-of-a-sparse-linear-model-using-ell_1-penalized-hubers-m-estimator
AUTHORS: Aram Dalalyan, Philip Thompson
HIGHLIGHT: We study the problem of estimating a p -dimensional s -sparse vector in a linear model with Gaussian design.

1183, TITLE: Learning nonlinear level sets for dimensionality reduction in function approximation
<https://papers.nips.cc/paper/9478-learning-nonlinear-level-sets-for-dimensionality-reduction-in-function-approximation>
AUTHORS: Guannan Zhang, Jiabin Zhang, Jacob Hinkle
HIGHLIGHT: We developed a Nonlinear Level-set Learning (NLL) method for dimensionality reduction in high-dimensional function approximation with small data.

1184, TITLE: Assessing Social and Intersectional Biases in Contextualized Word Representations
<https://papers.nips.cc/paper/9479-assessing-social-and-intersectional-biases-in-contextualized-word-representations>
AUTHORS: Yi Chern Tan, L. Elisa Celis
HIGHLIGHT: In this paper, we analyze the extent to which state-of-the-art models for contextual word representations, such as BERT and GPT-2, encode biases with respect to gender, race, and intersectional identities.

1185, TITLE: Online Convex Matrix Factorization with Representative Regions
<https://papers.nips.cc/paper/9480-online-convex-matrix-factorization-with-representative-regions>
AUTHORS: Jianhao Peng, Olgica Milenkovic, Abhishek Agarwal
HIGHLIGHT: We address both problems by proposing the first online convex MF algorithm that maintains a collection of constant-size sets of representative data samples needed for interpreting each of the basis (Ding et al., 2010) and has the same almost sure convergence guarantees as the online learning algorithm of Mairal et al., 2010.

- 1186, TITLE: Self-supervised GAN: Analysis and Improvement with Multi-class Minimax Game
<https://papers.nips.cc/paper/9481-self-supervised-gan-analysis-and-improvement-with-multi-class-minimax-game>
AUTHORS: Ngoc-Trung Tran, Viet-Hung Tran, Bao-Ngoc Nguyen, Linxiao Yang, Ngai-Man (Man) Cheung
HIGHLIGHT: In this work, we perform an in-depth analysis to understand how SS tasks interact with learning of generator.
- 1187, TITLE: Extreme Classification in Log Memory using Count-Min Sketch: A Case Study of Amazon Search with 50M Products
<https://papers.nips.cc/paper/9482-extreme-classification-in-log-memory-using-count-min-sketch-a-case-study-of-amazon-search-with-50m-products>
AUTHORS: Tharun Kumar Reddy Medini, Qixuan Huang, Yiqiu Wang, Vijai Mohan, Anshumali Shrivastava
HIGHLIGHT: To alleviate this problem, we present Merged-Average Classifiers via Hashing (MACH), a generic \mathcal{K} -classification algorithm where memory provably scales at $\mathcal{O}(\log K)$ without any assumption on the relation between classes.
- 1188, TITLE: A Fourier Perspective on Model Robustness in Computer Vision
<https://papers.nips.cc/paper/9483-a-fourier-perspective-on-model-robustness-in-computer-vision>
AUTHORS: Dong Yin, Raphael Gontijo Lopes, Jon Shlens, Ekin Dogus Cubuk, Justin Gilmer
HIGHLIGHT: We find that both methods improve robustness to corruptions that are concentrated in the high frequency domain while reducing robustness to corruptions that are concentrated in the low frequency domain.
- 1189, TITLE: The continuous Bernoulli: fixing a pervasive error in variational autoencoders
<https://papers.nips.cc/paper/9484-the-continuous-bernoulli-fixing-a-pervasive-error-in-variational-autoencoders>
AUTHORS: Gabriel Loaiza-Ganem, John P. Cunningham
HIGHLIGHT: We introduce and fully characterize a new $[0,1]$ -supported, single parameter distribution: the continuous Bernoulli, which patches this pervasive bug in VAE.
- 1190, TITLE: Privacy Amplification by Mixing and Diffusion Mechanisms
<https://papers.nips.cc/paper/9485-privacy-amplification-by-mixing-and-diffusion-mechanisms>
AUTHORS: Borja Balle, Gilles Barthe, Marco Gaboardi, Joseph Geumlek
HIGHLIGHT: In this paper we investigate under what conditions stochastic post-processing can amplify the privacy of a mechanism.
- 1191, TITLE: Variance Reduction in Bipartite Experiments through Correlation Clustering
<https://papers.nips.cc/paper/9486-variance-reduction-in-bipartite-experiments-through-correlation-clustering>
AUTHORS: Jean Pouget-Abadie, Kevin Aydin, Warren Schudy, Kay Brodersen, Vahab Mirrokni
HIGHLIGHT: This paper introduces a novel clustering objective and a corresponding algorithm that partitions a bipartite graph so as to maximize the statistical power of a bipartite experiment on that graph.
- 1192, TITLE: Gossip-based Actor-Learner Architectures for Deep Reinforcement Learning
<https://papers.nips.cc/paper/9487-gossip-based-actor-learner-architectures-for-deep-reinforcement-learning>
AUTHORS: Mahmoud ("Mido") Assran, Joshua Romoff, Nicolas Ballas, Joelle Pineau, Mike Rabbat
HIGHLIGHT: In this work, we propose Gossip-based Actor-Learner Architectures (GALA) where several actor-learners (such as A2C agents) are organized in a peer-to-peer communication topology, and exchange information through asynchronous gossip in order to take advantage of a large number of distributed simulators.
- 1193, TITLE: Metalearned Neural Memory
<https://papers.nips.cc/paper/9488-metalearned-neural-memory>
AUTHORS: Tsendsuren Munkhdalai, Alessandro Sordani, TONG WANG, Adam Trischler
HIGHLIGHT: We augment recurrent neural networks with an external memory mechanism that builds upon recent progress in metalearning.
- 1194, TITLE: Learning Multiple Markov Chains via Adaptive Allocation
<https://papers.nips.cc/paper/9489-learning-multiple-markov-chains-via-adaptive-allocation>
AUTHORS: Mohammad Sadegh Talebi, Odalric-Ambrym Maillard
HIGHLIGHT: We present a novel learning algorithm that efficiently balances *exploration* and *exploitation* intrinsic to this problem, without any prior knowledge of the chains.
- 1195, TITLE: Diffusion Improves Graph Learning
<https://papers.nips.cc/paper/9490-diffusion-improves-graph-learning>
AUTHORS: Johannes Klicpera, Stefan Weis, Wenzel, Stephan G. Schödl

HIGHLIGHT: In this work, we remove the restriction of using only the direct neighbors by introducing a powerful, yet spatially localized graph convolution: Graph diffusion convolution (GDC).

1196, **TITLE:** Deep Random Splines for Point Process Intensity Estimation of Neural Population Data
<https://papers.nips.cc/paper/9491-deep-random-splines-for-point-process-intensity-estimation-of-neural-population-data>
AUTHORS: Gabriel Loaiza-Ganem, Sean Perkins, Karen Schroeder, Mark Churchland, John P. Cunningham
HIGHLIGHT: Here we propose Deep Random Splines, a flexible class of random functions obtained by transforming Gaussian noise through a deep neural network whose output are the parameters of a spline.

1197, **TITLE:** Variational Bayes under Model Misspecification
<https://papers.nips.cc/paper/9492-variational-bayes-under-model-misspecification>
AUTHORS: Yixin Wang, David Blei
HIGHLIGHT: In this work, we study VB under model misspecification.

1198, **TITLE:** Global Convergence of Gradient Descent for Deep Linear Residual Networks
<https://papers.nips.cc/paper/9493-global-convergence-of-gradient-descent-for-deep-linear-residual-networks>
AUTHORS: Lei Wu, Qingcan Wang, Chao Ma
HIGHLIGHT: We analyze the global convergence of gradient descent for deep linear residual networks by proposing a new initialization: zero-asymmetric (ZAS) initialization.

1199, **TITLE:** On Differentially Private Graph Sparsification and Applications
<https://papers.nips.cc/paper/9494-on-differentially-private-graph-sparsification-and-applications>
AUTHORS: Raman Arora, Jalaj Upadhyay
HIGHLIGHT: In this paper, we study private sparsification of graphs.

1200, **TITLE:** Manifold denoising by Nonlinear Robust Principal Component Analysis
<https://papers.nips.cc/paper/9495-manifold-denoising-by-nonlinear-robust-principal-component-analysis>
AUTHORS: He Lyu, Ningyu Sha, Shuyang Qin, Ming Yan, Yuying Xie, Rongrong Wang
HIGHLIGHT: We answer these two questions affirmatively by proposing and analyzing an optimization framework that separates the sparse component from the manifold under noisy data.

1201, **TITLE:** Near-Optimal Reinforcement Learning in Dynamic Treatment Regimes
<https://papers.nips.cc/paper/9496-near-optimal-reinforcement-learning-in-dynamic-treatment-regimes>
AUTHORS: Junzhe Zhang, Elias Bareinboim
HIGHLIGHT: In this paper, we investigate the online reinforcement learning (RL) problem for selecting optimal DTRs provided that observational data is available.

1202, **TITLE:** ODE2VAE: Deep generative second order ODEs with Bayesian neural networks
<https://papers.nips.cc/paper/9497-ode2vae-deep-generative-second-order-odes-with-bayesian-neural-networks>
AUTHORS: Gagatay Yildiz, Markus Heinonen, Harri Lahdesmaki
HIGHLIGHT: We present Ordinary Differential Equation Variational Auto-Encoder (ODE2VAE), a latent second order ODE model for high-dimensional sequential data.

1203, **TITLE:** Optimal Sampling and Clustering in the Stochastic Block Model
<https://papers.nips.cc/paper/9498-optimal-sampling-and-clustering-in-the-stochastic-block-model>
AUTHORS: Se-Young Yun, Alexandre Proutiere
HIGHLIGHT: This paper investigates the design of joint adaptive sampling and clustering algorithms in networks whose structure follows the celebrated Stochastic Block Model (SBM).

1204, **TITLE:** Recurrent Kernel Networks
<https://papers.nips.cc/paper/9499-recurrent-kernel-networks>
AUTHORS: Dexiong Chen, Laurent Jacob, Julien Mairal
HIGHLIGHT: In this paper, we revisit this link by generalizing convolutional kernel networks---originally related to a relaxation of the mismatch kernel---to model gaps in sequences.

1205, **TITLE:** Cold Case: The Lost MNIST Digits
<https://papers.nips.cc/paper/9500-cold-case-the-lost-mnist-digits>
AUTHORS: Chhavi Yadav, Leon Bottou
HIGHLIGHT: We propose a reconstruction that is accurate enough to serve as a replacement for the MNIST dataset, with insignificant changes in accuracy.

- 1206, TITLE: Hierarchical Optimal Transport for Multimodal Distribution Alignment
<https://papers.nips.cc/paper/9501-hierarchical-optimal-transport-for-multimodal-distribution-alignment>
AUTHORS: John Lee, Max Dabagia, Eva Dyer, Christopher Rozell
HIGHLIGHT: To solve this numerically, we propose a distributed ADMM algorithm that also exploits the Sinkhorn distance, thus it has an efficient computational complexity that scales quadratically with the size of the largest cluster.
- 1207, TITLE: Exploration via Hindsight Goal Generation
<https://papers.nips.cc/paper/9502-exploration-via-hindsight-goal-generation>
AUTHORS: Zhizhou Ren, Kefan Dong, Yuan Zhou, Qiang Liu, Jian Peng
HIGHLIGHT: In this paper, we introduce Hindsight Goal Generation (HGG), a novel algorithmic framework that generates valuable hindsight goals which are easy for an agent to achieve in the short term and are also potential for guiding the agent to reach the actual goal in the long term.
- 1208, TITLE: Shaping Belief States with Generative Environment Models for RL
<https://papers.nips.cc/paper/9503-shaping-belief-states-with-generative-environment-models-for-rl>
AUTHORS: Karol Gregor, Danilo Jimenez Rezende, Frederic Besse, Yan Wu, Hamza Merzic, Aaron van den Oord
HIGHLIGHT: We propose a way to efficiently train expressive generative models in complex environments.
- 1209, TITLE: Globally Optimal Learning for Structured Elliptical Losses
<https://papers.nips.cc/paper/9504-globally-optimal-learning-for-structured-elliptical-losses>
AUTHORS: Yoav Wald, Nofar Noy, Gal Elidan, Ami Wiesel
HIGHLIGHT: In this work, we analyze robust alternatives.
- 1210, TITLE: Object landmark discovery through unsupervised adaptation
<https://papers.nips.cc/paper/9505-object-landmark-discovery-through-unsupervised-adaptation>
AUTHORS: Enrique Sanchez, Georgios Tzimiopoulos
HIGHLIGHT: This paper proposes a method to ease the unsupervised learning of object landmark detectors.
- 1211, TITLE: Specific and Shared Causal Relation Modeling and Mechanism-Based Clustering
<https://papers.nips.cc/paper/9506-specific-and-shared-causal-relation-modeling-and-mechanism-based-clustering>
AUTHORS: Biwei Huang, Kun Zhang, Pengtao Xie, Mingming Gong, Eric P. Xing, Clark Glymour
HIGHLIGHT: In this paper, we develop a unified framework for causal discovery and mechanism-based group identification.
- 1212, TITLE: Search-Guided, Lightly-Supervised Training of Structured Prediction Energy Networks
<https://papers.nips.cc/paper/9507-search-guided-lightly-supervised-training-of-structured-prediction-energy-networks>
AUTHORS: Amirmohammad Rooshenas, Dongxu Zhang, Gopal Sharma, Andrew McCallum
HIGHLIGHT: In this paper, we instead use efficient truncated randomized search in this reward function to train structured prediction energy networks (SPENs), which provide efficient test-time inference using gradient-based search on a smooth, learned representation of the score landscape, and have previously yielded state-of-the-art results in structured prediction.
- 1213, TITLE: Accelerating Rescaled Gradient Descent: Fast Optimization of Smooth Functions
<https://papers.nips.cc/paper/9508-accelerating-rescaled-gradient-descent-fast-optimization-of-smooth-functions>
AUTHORS: Ashia C. Wilson, Lester Mackey, Andre Wibisono
HIGHLIGHT: We present a family of algorithms, called descent algorithms, for optimizing convex and non-convex functions.
- 1214, TITLE: RUDDER: Return Decomposition for Delayed Rewards
<https://papers.nips.cc/paper/9509-rudder-return-decomposition-for-delayed-rewards>
AUTHORS: Jose A. Arjona-Medina, Michael Gillhofer, Michael Widrich, Thomas Unterthiner, Johannes Brandstetter, Sepp Hochreiter
HIGHLIGHT: We propose RUDDER, a novel reinforcement learning approach for delayed rewards in finite Markov decision processes (MDPs).
- 1215, TITLE: Graph Normalizing Flows
<https://papers.nips.cc/paper/9510-graph-normalizing-flows>
AUTHORS: Jenny Liu, Aviral Kumar, Jimmy Ba, Jamie Kiros, Kevin Swersky
HIGHLIGHT: We introduce graph normalizing flows: a new, reversible graph neural network model for prediction and generation.

- 1216, TITLE: Explanations can be manipulated and geometry is to blame
<https://papers.nips.cc/paper/9511-explanations-can-be-manipulated-and-geometry-is-to-blame>
AUTHORS: Ann-Kathrin Dombrowski, Maximilian Alber, Christopher Anders, Marcel Ackermann, Klaus-Robert Müller, Pan Kessel
HIGHLIGHT: In this paper, we demonstrate a property of explanation methods which is disconcerting for both of these purposes.
- 1217, TITLE: Communication trade-offs for Local-SGD with large step size
<https://papers.nips.cc/paper/9512-communication-trade-offs-for-local-sgd-with-large-step-size>
AUTHORS: Aymeric Dieuleveut, Kumar Kshitij Patel
HIGHLIGHT: We propose a non-asymptotic error analysis, which enables comparison to \emph{one-shot averaging} i.e., a single communication round among independent workers, and \emph{mini-batch averaging} i.e., communicating at every step.
- 1218, TITLE: Non-normal Recurrent Neural Network (nnRNN): learning long time dependencies while improving expressivity with transient dynamics
<https://papers.nips.cc/paper/9513-non-normal-recurrent-neural-network-nnrmn-learning-long-time-dependencies-while-improving-expressivity-with-transient-dynamics>
AUTHORS: Giancarlo Kerg, Kyle Goyette, Maximilian Puelma Touzel, Gauthier Gidel, Eugene Vorontsov, Yoshua Bengio, Guillaume Lajoie
HIGHLIGHT: We propose a novel connectivity structure based on the Schur decomposition and a splitting of the Schur form into normal and non-normal parts.
- 1219, TITLE: No-Regret Learning in Unknown Games with Correlated Payoffs
<https://papers.nips.cc/paper/9514-no-regret-learning-in-unknown-games-with-correlated-payoffs>
AUTHORS: Pier Giuseppe Sessa, Ilija Bogunovic, Maryam Kamgarpour, Andreas Krause
HIGHLIGHT: In this paper, we consider a natural model where, besides a noisy measurement of the obtained reward, the player can also observe the opponents' actions.
- 1220, TITLE: Alleviating Label Switching with Optimal Transport
<https://papers.nips.cc/paper/9515-alleviating-label-switching-with-optimal-transport>
AUTHORS: Pierre Monteleiller, Sebastian Claiici, Edward Chien, Farzaneh Mirzazadeh, Justin M. Solomon, Mikhail Yurochkin
HIGHLIGHT: We propose a resolution to label switching that leverages machinery from optimal transport.
- 1221, TITLE: Paraphrase Generation with Latent Bag of Words
<https://papers.nips.cc/paper/9516-paraphrase-generation-with-latent-bag-of-words>
AUTHORS: Yao Fu, Yansong Feng, John P. Cunningham
HIGHLIGHT: Inspired by variational autoencoders with discrete latent structures, in this work, we propose a latent bag of words (BOW) model for paraphrase generation.
- 1222, TITLE: An Algorithmic Framework For Differentially Private Data Analysis on Trusted Processors
<https://papers.nips.cc/paper/9517-an-algorithmic-framework-for-differentially-private-data-analysis-on-trusted-processors>
AUTHORS: Joshua Allen, Bolin Ding, Janardhan Kulkarni, Harsha Nori, Olga Ohrimenko, Sergey Yekhanin
HIGHLIGHT: In this work, we propose a framework based on trusted processors and a new definition of differential privacy called Oblivious Differential Privacy, which combines the best of both local and global models.
- 1223, TITLE: Compacting, Picking and Growing for Unforgetting Continual Learning
<https://papers.nips.cc/paper/9518-compacting-picking-and-growing-for-unforgetting-continual-learning>
AUTHORS: Ching-Yi Hung, Cheng-Hao Tu, Cheng-En Wu, Chien-Hung Chen, Yi-Ming Chan, Chu-Song Chen
HIGHLIGHT: In this paper, we propose a simple but effective approach to continual deep learning.
- 1224, TITLE: Approximating Interactive Human Evaluation with Self-Play for Open-Domain Dialog Systems
<https://papers.nips.cc/paper/9519-approximating-interactive-human-evaluation-with-self-play-for-open-domain-dialog-systems>
AUTHORS: Asma Ghandeharion, Judy Hanwen Shen, Natasha Jaques, Craig Ferguson, Noah Jones, Agata Lapedriza, Rosalind Picard
HIGHLIGHT: In particular, we propose a self-play scenario where the dialog system talks to itself and we calculate a combination of proxies such as sentiment and semantic coherence on the conversation trajectory.
- 1225, TITLE: A New Distribution on the Simplex with Auto-Encoding Applications
<https://papers.nips.cc/paper/9520-a-new-distribution-on-the-simplex-with-auto-encoding-applications>
AUTHORS: Andrew Stim, Tony Jebara, David Knowles

HIGHLIGHT: We construct a new distribution for the simplex using the Kumaraswamy distribution and an ordered stick-breaking process.

1226, **TITLE:** AutoPrune: Automatic Network Pruning by Regularizing Auxiliary Parameters
<https://papers.nips.cc/paper/9521-autoprune-automatic-network-pruning-by-regularizing-auxiliary-parameters>

AUTHORS: XIA XIAO, Zigeng Wang, Sanguthevar Rajasekaran

HIGHLIGHT: To build a better generalized and easy-to-use pruning method, we propose AutoPrune, which prunes the network through optimizing a set of trainable auxiliary parameters instead of original weights.

1227, **TITLE:** A neurally plausible model learns successor representations in partially observable environments
<https://papers.nips.cc/paper/9522-a-neurally-plausible-model-learns-successor-representations-in-partially-observable-environments>

AUTHORS: Eszter V?rtes, Maneesh Sahani

HIGHLIGHT: Here, we introduce a neurally plausible model using $\text{distributonal successor features}$, which builds on the distributed distributional code for the representation and computation of uncertainty, and which allows for efficient value function computation in partially observed environments via the successor representation.

1228, **TITLE:** Learning about an exponential amount of conditional distributions

<https://papers.nips.cc/paper/9523-learning-about-an-exponential-amount-of-conditional-distributions>

AUTHORS: Mohamed Belghazi, Maxime Oquab, David Lopez-Paz

HIGHLIGHT: We introduce the Neural Conditioner (NC), a self-supervised machine able to learn about all the conditional distributions of a random vector X .

1229, **TITLE:** Towards modular and programmable architecture search

<https://papers.nips.cc/paper/9524-towards-modular-and-programmable-architecture-search>

AUTHORS: Renato Negrinho, Matthew Gormley, Geoffrey J. Gordon, Darshan Patil, Nghia Le, Daniel Ferreira

HIGHLIGHT: In this work, we propose a formal language for encoding search spaces over general computational graphs.

1230, **TITLE:** Towards Hardware-Aware Tractable Learning of Probabilistic Models

<https://papers.nips.cc/paper/9525-towards-hardware-aware-tractable-learning-of-probabilistic-models>

AUTHORS: Laura I. Galindez Olascoaga, Wannes Meert, Nimish Shah, Marian Verhelst, Guy Van den Broeck

HIGHLIGHT: We propose a novel resource-aware cost metric that takes into consideration the hardware's properties in determining whether the inference task can be efficiently deployed.

1231, **TITLE:** On Robustness to Adversarial Examples and Polynomial Optimization

<https://papers.nips.cc/paper/9526-on-robustness-to-adversarial-examples-and-polynomial-optimization>

AUTHORS: Pranjali Awasthi, Abhratanu Dutta, Aravindan Vijayaraghavan

HIGHLIGHT: The main contribution of this work is to exhibit a strong connection between achieving robustness to adversarial examples, and a rich class of polynomial optimization problems, thereby making progress on the above questions.

1232, **TITLE:** Rand-NSG: Fast Accurate Billion-point Nearest Neighbor Search on a Single Node

<https://papers.nips.cc/paper/9527-rand-nsg-fast-accurate-billion-point-nearest-neighbor-search-on-a-single-node>

AUTHORS: Suhas Jayaram Subramanya, Fnu Devvrit, Harsha Vardhan Simhadri, Ravishankar Krishnawamy, Rohan Kadekodi

HIGHLIGHT: We present a new graph-based indexing and search system called DiskANN that can index, store, and search a billion point database on a single workstation with just 64GB RAM and an inexpensive solid-state drive (SSD).

1233, **TITLE:** A Solvable High-Dimensional Model of GAN

<https://papers.nips.cc/paper/9528-a-solvable-high-dimensional-model-of-gan>

AUTHORS: Chuang Wang, Hong Hu, Yue Lu

HIGHLIGHT: We present a theoretical analysis of the training process for a single-layer GAN fed by high-dimensional input data.

1234, **TITLE:** Using Embeddings to Correct for Unobserved Confounding in Networks

<https://papers.nips.cc/paper/9529-using-embeddings-to-correct-for-unobserved-confounding-in-networks>

AUTHORS: Victor Veitch, Yixin Wang, David Blei

HIGHLIGHT: We consider causal inference in the presence of unobserved confounding.

1235, **TITLE:** MonoForest framework for tree ensemble analysis

<https://papers.nips.cc/paper/9530-monoforest-framework-for-tree-ensemble-analysis>

AUTHORS: Igor Kuralenok, Vasilii Ershov, Igor Labutin

HIGHLIGHT: In this work, we introduce a new decision tree ensemble representation framework: instead of using a graph model we transform each tree into a well-known polynomial form.

1236, **TITLE:** Bayesian Optimization under Heavy-tailed Payoffs
<https://papers.nips.cc/paper/9531-bayesian-optimization-under-heavy-tailed-payoffs>

AUTHORS: Sayak Ray Chowdhury, Aditya Gopalan

HIGHLIGHT: We consider black box optimization of an unknown function in the nonparametric Gaussian process setting when the noise in the observed function values can be heavy tailed.

1237, **TITLE:** Combining Generative and Discriminative Models for Hybrid Inference
<https://papers.nips.cc/paper/9532-combining-generative-and-discriminative-models-for-hybrid-inference>

AUTHORS: Victor Garcia Satorras, Max Welling, Zeynep Akata

HIGHLIGHT: In this work we propose a hybrid model that combines graphical inference with a learned inverse model, which we structure as in a graph neural network, while the iterative algorithm as a whole is formulated as a recurrent neural network.

1238, **TITLE:** A Graph Theoretic Additive Approximation of Optimal Transport
<https://papers.nips.cc/paper/9533-a-graph-theoretic-additive-approximation-of-optimal-transport>

AUTHORS: Nathaniel Lahn, Deepika Mulchandani, Sharath Raghvendra

HIGHLIGHT: We present an adaptation of the classical graph algorithm of Gabow and Tarjan and provide a novel analysis of this algorithm that bounds its execution time by $\mathcal{O}\left(\frac{n^2 C}{\delta} + \frac{n C^2}{\delta^2}\right)$.

1239, **TITLE:** Adversarial Robustness through Local Linearization

<https://papers.nips.cc/paper/9534-adversarial-robustness-through-local-linearization>

AUTHORS: Chongli Qin, James Martens, Sven Gowal, Dilip Krishnan, Krishnamurthy Dvijotham, Alhussein Fawzi, Soham De, Robert Stanforth, Pushmeet Kohli

HIGHLIGHT: In this work, we introduce a novel regularizer that encourages the loss to behave linearly in the vicinity of the training data, thereby penalizing gradient obfuscation while encouraging robustness.

1240, **TITLE:** Sampled Softmax with Random Fourier Features

<https://papers.nips.cc/paper/9535-sampled-softmax-with-random-fourier-features>

AUTHORS: Ankit Singh Rawat, Jiecao Chen, Felix Xinnan X. Yu, Ananda Theertha Suresh, Sanjiv Kumar

HIGHLIGHT: Motivated by our analysis and the work on kernel-based sampling, we propose the Random Fourier Softmax (RF-softmax) method that utilizes the powerful Random Fourier Features to enable more efficient and accurate sampling from an approximate softmax distribution.

1241, **TITLE:** Semi-flat minima and saddle points by embedding neural networks to overparameterization

<https://papers.nips.cc/paper/9536-semi-flat-minima-and-saddle-points-by-embedding-neural-networks-to-overparameterization>

AUTHORS: Kenji Fukumizu, Shoichiro Yamaguchi, Yoh-ichi Mototake, Mirai Tanaka

HIGHLIGHT: We consider three basic methods for embedding a network into a wider one with more hidden units, and discuss whether a minimum point of the narrower network gives a minimum or saddle point of the wider one.

1242, **TITLE:** Learning Fairness in Multi-Agent Systems

<https://papers.nips.cc/paper/9537-learning-fairness-in-multi-agent-systems>

AUTHORS: Jiechuan Jiang, Zongqing Lu

HIGHLIGHT: To tackle these difficulties, we propose FEN, a novel hierarchical reinforcement learning model.

1243, **TITLE:** Primal-Dual Block Generalized Frank-Wolfe

<https://papers.nips.cc/paper/9538-primal-dual-block-generalized-frank-wolfe>

AUTHORS: Qi Lei, JIACHENG ZHUO, Constantine Caramanis, Inderjit S. Dhillon, Alexandros G. Dimakis

HIGHLIGHT: We propose a generalized variant of Frank-Wolfe algorithm for solving a class of sparse/low-rank optimization problems.

1244, **TITLE:** GOT: An Optimal Transport framework for Graph comparison

<https://papers.nips.cc/paper/9539-got-an-optimal-transport-framework-for-graph-comparison>

AUTHORS: Hermina Petric Maretic, Mireille El Gheche, Giovanni Chierchia, Pascal Frossard

HIGHLIGHT: We present a novel framework based on optimal transport for the challenging problem of comparing graphs.

1245, **TITLE:** On Mixup Training: Improved Calibration and Predictive Uncertainty for Deep Neural Networks

<https://papers.nips.cc/paper/9540-on-mixup-training-improved-calibration-and-predictive-uncertainty-for-deep-neural-networks>

AUTHORS: Sunil Thulasidasan, Gopinath Chennupati, Jeff A. Bilmes, Tanmoy Bhattacharya, Sarah Michalak

HIGHLIGHT: In this work, we discuss a hitherto untouched aspect of mixup training -- the calibration and predictive uncertainty of models trained with mixup.

1246, **TITLE:** Complexity of Highly Parallel Non-Smooth Convex Optimization
<https://papers.nips.cc/paper/9541-complexity-of-highly-parallel-non-smooth-convex-optimization>
AUTHORS: Sebastien Bubeck, Qijia Jiang, Yin-Tat Lee, Yuanzhi Li, Aaron Sidford
HIGHLIGHT: Namely we consider optimization algorithms interacting with a highly parallel gradient oracle, that is one that can answer poly(d) gradient queries in parallel.

1247, **TITLE:** Inverting Deep Generative models, One layer at a time
<https://papers.nips.cc/paper/9542-inverting-deep-generative-models-one-layer-at-a-time>
AUTHORS: Qi Lei, Ajil Jalal, Inderjit S. Dhillon, Alexandros G. Dimakis
HIGHLIGHT: In this paper we obtain several novel theoretical results for the inversion problem.

1248, **TITLE:** Calculating Optimistic Likelihoods Using (Geodesically) Convex Optimization
<https://papers.nips.cc/paper/9543-calculating-optimistic-likelihoods-using-geodesically-convex-optimization>
AUTHORS: Viet Anh Nguyen, Soroosh Shafieezadeh Abadeh, Man-Chung Yue, Daniel Kuhn, Wolfram Wiesemann
HIGHLIGHT: We thus propose to replace each nominal distribution with an ambiguity set containing all distributions in its vicinity and to evaluate an optimistic likelihood, that is, the maximum of the likelihood over all distributions in the ambiguity set.

1249, **TITLE:** The Implicit Metropolis-Hastings Algorithm
<https://papers.nips.cc/paper/9544-the-implicit-metropolis-hastings-algorithm>
AUTHORS: Kirill Neklyudov, Evgenii Egorov, Dmitry P. Vetrov
HIGHLIGHT: For any implicit probabilistic model and a target distribution represented by a set of samples, implicit Metropolis-Hastings operates by learning a discriminator to estimate the density-ratio and then generating a chain of samples.

1250, **TITLE:** An Inexact Augmented Lagrangian Framework for Nonconvex Optimization with Nonlinear Constraints
<https://papers.nips.cc/paper/9545-an-inexact-augmented-lagrangian-framework-for-nonconvex-optimization-with-nonlinear-constraints>
AUTHORS: Mehmet Fatih Sahin, Armin eftekhari, Ahmet Alacaoglu, Fabian Latorre, Volkan Cevher
HIGHLIGHT: We propose a practical inexact augmented Lagrangian method (iALM) for nonconvex problems with nonlinear constraints.

1251, **TITLE:** Generalization in Reinforcement Learning with Selective Noise Injection and Information Bottleneck
<https://papers.nips.cc/paper/9546-generalization-in-reinforcement-learning-with-selective-noise-injection-and-information-bottleneck>
AUTHORS: Maximilian Igl, Kamil Ciosek, Yingzhen Li, Sebastian Tschiatschek, Cheng Zhang, Sam Devlin, Katja Hofmann
HIGHLIGHT: We discuss those differences and propose modifications to existing regularization techniques in order to better adapt them to RL.

1252, **TITLE:** Can you trust your model's uncertainty? Evaluating predictive uncertainty under dataset shift
<https://papers.nips.cc/paper/9547-can-you-trust-your-models-uncertainty-evaluating-predictive-uncertainty-under-dataset-shift>
AUTHORS: Jasper Snoek, Yaniv Ovadia, Emily Fertig, Balaji Lakshminarayanan, Sebastian Nowozin, D. Sculley, Joshua Dillon, Jie Ren, Zachary Nado
HIGHLIGHT: We present a large-scale benchmark of existing state-of-the-art methods on classification problems and investigate the effect of dataset shift on accuracy and calibration.

1253, **TITLE:** Accurate Layerwise Interpretable Competence Estimation
<https://papers.nips.cc/paper/9548-accurate-layerwise-interpretable-competence-estimation>
AUTHORS: Vickram Rajendran, William LeVine
HIGHLIGHT: In this paper, we seek to examine, understand, and predict the pointwise competence of classification models.

1254, **TITLE:** A New Perspective on Pool-Based Active Classification and False-Discovery Control
<https://papers.nips.cc/paper/9549-a-new-perspective-on-pool-based-active-classification-and-false-discovery-control>
AUTHORS: Lalit Jain, Kevin G. Jamieson
HIGHLIGHT: In this paper, we provide the first provably sample efficient adaptive algorithm for this problem.

1255, **TITLE:** Defending Neural Backdoors via Generative Distribution Modeling
<https://papers.nips.cc/paper/9550-defending-neural-backdoors-via-generative-distribution-modeling>
AUTHORS: Ximing Qiao, Yukun Yang, Hai Li

HIGHLIGHT: In the work, we explore the space formed by the pixel values of all possible backdoor triggers.

1256, TITLE: Are Sixteen Heads Really Better than One?
<https://papers.nips.cc/paper/9551-are-sixteen-heads-really-better-than-one>

AUTHORS: Paul Michel, Omer Levy, Graham Neubig

HIGHLIGHT: However we observe that, in practice, a large proportion of attention heads can be removed at test time without significantly impacting performance, and that some layers can even be reduced to a single head.

1257, TITLE: Multi-resolution Multi-task Gaussian Processes
<https://papers.nips.cc/paper/9552-multi-resolution-multi-task-gaussian-processes>

AUTHORS: Oliver Hamelijnck, Theodoros Damoulas, Kangrui Wang, Mark Girolami

HIGHLIGHT: We offer a multi-resolution multi-task (MRGP) framework that allows for both inter-task and intra-task multi-resolution and multi-fidelity.

1258, TITLE: Variational Bayesian Optimal Experimental Design
<https://papers.nips.cc/paper/9553-variational-bayesian-optimal-experimental-design>

AUTHORS: Adam Foster, Martin Jankowiak, Elias Bingham, Paul Horsfall, Yee Whye Teh, Thomas Rainforth, Noah Goodman

HIGHLIGHT: To address this, we introduce several classes of fast EIG estimators by building on ideas from amortized variational inference.

1259, TITLE: Universal Approximation of Input-Output Maps by Temporal Convolutional Nets
<https://papers.nips.cc/paper/9554-universal-approximation-of-input-output-maps-by-temporal-convolutional-nets>

AUTHORS: Joshua Hanson, Maxim Raginsky

HIGHLIGHT: We prove that TCNs can approximate a large class of input-output maps having approximately finite memory to arbitrary error tolerance.

1260, TITLE: Provable Certificates for Adversarial Examples: Fitting a Ball in the Union of Polytopes
<https://papers.nips.cc/paper/9555-provable-certificates-for-adversarial-examples-fitting-a-ball-in-the-union-of-polytopes>

AUTHORS: Matt Jordan, Justin Lewis, Alexandros G. Dimakis

HIGHLIGHT: We propose a novel method for computing exact pointwise robustness of deep neural networks for all convex l_p norms.

1261, TITLE: Reinforcement Learning with Convex Constraints
<https://papers.nips.cc/paper/9556-reinforcement-learning-with-convex-constraints>

AUTHORS: Sobhan Miryoosefi, Kiant? Brantley, Hal Daume III, Miro Dudik, Robert E. Schapire

HIGHLIGHT: In this paper, we propose an algorithmic scheme that can handle a wide class of constraints in RL tasks: specifically, any constraints that require expected values of some vector measurements (such as the use of an action) to lie in a convex set.

1262, TITLE: User-Specified Local Differential Privacy in Unconstrained Adaptive Online Learning
<https://papers.nips.cc/paper/9557-user-specified-local-differential-privacy-in-unconstrained-adaptive-online-learning>

AUTHORS: Dirk van der Hoeven

HIGHLIGHT: In this paper we generalize this approach by allowing the provider of the data to choose the distribution of the noise without disclosing any parameters of the distribution to the learner, under the constraint that the distribution is symmetrical.

1263, TITLE: Stochastic Bandits with Context Distributions
<https://papers.nips.cc/paper/9558-stochastic-bandits-with-context-distributions>

AUTHORS: Johannes Kirschner, Andreas Krause

HIGHLIGHT: We introduce a stochastic contextual bandit model where at each time step the environment chooses a distribution over a context set and samples the context from this distribution.

1264, TITLE: Inducing brain-relevant bias in natural language processing models
<https://papers.nips.cc/paper/9559-inducing-brain-relevant-bias-in-natural-language-processing-models>

AUTHORS: Dan Schwartz, Mariya Toneva, Leila Wehbe

HIGHLIGHT: We demonstrate that a version of BERT, a recently introduced and powerful language model, can improve the prediction of brain activity after fine-tuning.

1265, TITLE: Using a Logarithmic Mapping to Enable Lower Discount Factors in Reinforcement Learning
<https://papers.nips.cc/paper/9560-using-a-logarithmic-mapping-to-enable-lower-discount-factors-in-reinforcement-learning>

AUTHORS: Harm Van Seijen, Mehdi Fatemi, Arash Tavakoli
HIGHLIGHT: We propose an alternative hypothesis that identifies the size-difference of the action-gap across the state-space as the primary cause.

1266, TITLE: Recovering Bandits
<https://papers.nips.cc/paper/9561-recovering-bandits>
AUTHORS: Ciara Pike-Burke, Steffen Grunewalder
HIGHLIGHT: In this work, we explore the use of Gaussian processes to tackle the estimation and planing problem.

1267, TITLE: Computing Linear Restrictions of Neural Networks
<https://papers.nips.cc/paper/9562-computing-linear-restrictions-of-neural-networks>
AUTHORS: Matthew Sotoudeh, Aditya V. Thakur
HIGHLIGHT: We present an efficient algorithm for computing ExactLine for networks that use ReLU, MaxPool, batch normalization, fully-connected, convolutional, and other layers, along with several applications.

1268, TITLE: Learning Positive Functions with Pseudo Mirror Descent
<https://papers.nips.cc/paper/9563-learning-positive-functions-with-pseudo-mirror-descent>
AUTHORS: Yingxiang Yang, Haoxiang Wang, Negar Kiyavash, Niao He
HIGHLIGHT: In this paper, we propose a novel algorithm, pseudo mirror descent, that performs efficient estimation of positive functions within a Hilbert space without expensive projections.

1269, TITLE: Correlation Priors for Reinforcement Learning
<https://papers.nips.cc/paper/9564-correlation-priors-for-reinforcement-learning>
AUTHORS: Bastian Alt, Adrian ?o?ic, Heinz Koepl
HIGHLIGHT: In this work, we present a Bayesian learning framework based on Pólya-Gamma augmentation that enables an analogous reasoning in such cases.

1270, TITLE: Fast, Provably convergent IRLS Algorithm for p-norm Linear Regression
<https://papers.nips.cc/paper/9565-fast-provably-convergent-irls-algorithm-for-p-norm-linear-regression>
AUTHORS: Deeksha Adil, Richard Peng, Sushant Sachdeva
HIGHLIGHT: We propose p-IRLS, the first IRLS algorithm that provably converges geometrically for any $p \in [2, \infty)$.

1271, TITLE: A Similarity-preserving Network Trained on Transformed Images Recapitulates Salient Features of the Fly Motion Detection Circuit
<https://papers.nips.cc/paper/9566-a-similarity-preserving-network-trained-on-transformed-images-recapitulates-salient-features-of-the-fly-motion-detection-circuit>
AUTHORS: Yanis Bahroun, Dmitri Chklovskii, Anirvan Sengupta
HIGHLIGHT: Here we propose a biologically plausible model of motion detection.

1272, TITLE: Differentially Private Covariance Estimation
<https://papers.nips.cc/paper/9567-differentially-private-covariance-estimation>
AUTHORS: Kareem Amin, Travis Dick, Alex Kulesza, Andres Munoz, Sergei Vassilvitskii
HIGHLIGHT: In this work we propose a new epsilon-differentially private algorithm for computing the covariance matrix of a dataset that addresses both of these limitations.

1273, TITLE: Outlier Detection and Robust PCA Using a Convex Measure of Innovation
<https://papers.nips.cc/paper/9568-outlier-detection-and-robust-pca-using-a-convex-measure-of-innovation>
AUTHORS: Mostafa Rahmani, Ping Li
HIGHLIGHT: This paper presents a provable and strong algorithm, termed Innovation Search (iSearch), to robust Principal Component Analysis (PCA) and outlier detection.

1274, TITLE: Integrating Markov processes with structural causal modeling enables counterfactual inference in complex systems
<https://papers.nips.cc/paper/9569-integrating-markov-processes-with-structural-causal-modeling-enables-counterfactual-inference-in-complex-systems>
AUTHORS: Robert Ness, Kaushal Paneri, Olga Vitek
HIGHLIGHT: This manuscript contributes a general and practical framework for casting a Markov process model of a system at equilibrium as a structural causal model, and carrying out counterfactual inference.

1275, TITLE: Are Disentangled Representations Helpful for Abstract Visual Reasoning?

- <https://papers.nips.cc/paper/9570-are-disentangled-representations-helpful-for-abstract-visual-reasoning>
AUTHORS: Sjoerd van Steenkiste, Francesco Locatello, J?rgen Schmidhuber, Olivier Bachem
HIGHLIGHT: In this paper, we conduct a large-scale study that investigates whether disentangled representations are more suitable for abstract reasoning tasks.
- 1276, TITLE: PowerSGD: Practical Low-Rank Gradient Compression for Distributed Optimization
<https://papers.nips.cc/paper/9571-powersgd-practical-low-rank-gradient-compression-for-distributed-optimization>
AUTHORS: Thijs Vogels, Sai Praneeth Karimireddy, Martin Jaggi
HIGHLIGHT: We propose a low-rank gradient compressor that can i) compress gradients rapidly, ii) efficiently aggregate the compressed gradients using all-reduce, and iii) achieve test performance on par with SGD.
- 1277, TITLE: Stochastic Frank-Wolfe for Composite Convex Minimization
<https://papers.nips.cc/paper/9572-stochastic-frank-wolfe-for-composite-convex-minimization>
AUTHORS: Francesco Locatello, Alp Yurtsever, Olivier Fercoq, Volkan Cevher
HIGHLIGHT: In this work, we propose the first conditional-gradient-type method for solving stochastic optimization problems under affine constraints.
- 1278, TITLE: Constraint-based Causal Structure Learning with Consistent Separating Sets
<https://papers.nips.cc/paper/9573-constraint-based-causal-structure-learning-with-consistent-separating-sets>
AUTHORS: Honghao Li, Vincent Cabeli, Nadir Sella, Herve Isambert
HIGHLIGHT: In this paper, we propose a simple modification of PC and PC-derived algorithms so as to ensure that all separating sets identified to remove dispensable edges are consistent with the final graph, thus enhancing the explainability of constraint-based methods.
- 1279, TITLE: Unsupervised Discovery of Temporal Structure in Noisy Data with Dynamical Components Analysis
<https://papers.nips.cc/paper/9574-unsupervised-discovery-of-temporal-structure-in-noisy-data-with-dynamical-components-analysis>
AUTHORS: David Clark, Jesse Livezey, Kristofer Bouchard
HIGHLIGHT: Combining these approaches, we introduce Dynamical Components Analysis (DCA), a linear dimensionality reduction method which discovers a subspace of high-dimensional time series data with maximal predictive information, defined as the mutual information between the past and future.
- 1280, TITLE: Sample Efficient Active Learning of Causal Trees
<https://papers.nips.cc/paper/9575-sample-efficient-active-learning-of-causal-trees>
AUTHORS: Kristjan Greenewald, Dmitriy Katz, Karthikeyan Shanmugam, Sara Magliacane, Murat Kocaoglu, Enric Boix Adsera, Guy Bresler
HIGHLIGHT: We propose an adaptive framework that determines the next intervention based on a Bayesian prior updated with the outcomes of previous experiments, focusing on the setting where observational data is cheap (assumed infinite) and interventional data is expensive.
- 1281, TITLE: Efficient Neural Architecture Transformation Search in Channel-Level for Object Detection
<https://papers.nips.cc/paper/9576-efficient-neural-architecture-transformation-search-in-channel-level-for-object-detection>
AUTHORS: Junran Peng, Ming Sun, ZHAO-XIANG ZHANG, Tieniu Tan, Junjie Yan
HIGHLIGHT: To overcome this obstacle, we introduce a practical neural architecture transformation search (NATS) algorithm for object detection in this paper.
- 1282, TITLE: Robust Attribution Regularization
<https://papers.nips.cc/paper/9577-robust-attribution-regularization>
AUTHORS: Jiefeng Chen, Xi Wu, Vaibhav Rastogi, Yingyu Liang, Somesh Jha
HIGHLIGHT: We propose training objectives in classic robust optimization models to achieve robust IG attributions.
- 1283, TITLE: Computational Mirrors: Blind Inverse Light Transport by Deep Matrix Factorization
<https://papers.nips.cc/paper/9578-computational-mirrors-blind-inverse-light-transport-by-deep-matrix-factorization>
AUTHORS: Miika Aittala, Prafull Sharma, Lukas Murmann, Adam Yedidia, Gregory Wornell, Bill Freeman, Fredo Durand
HIGHLIGHT: We solve this problem by factoring the observed video into a matrix product between the unknown hidden scene video and an unknown light transport matrix.
- 1284, TITLE: When to use parametric models in reinforcement learning?
<https://papers.nips.cc/paper/9579-when-to-use-parametric-models-in-reinforcement-learning>
AUTHORS: Hado P. van Hasselt, Matteo Hessel, John Aslanides
HIGHLIGHT: We examine the question of when and how parametric models are most useful in reinforcement learning.

- 1285, TITLE: General E(2)-Equivariant Steerable CNNs
<https://papers.nips.cc/paper/9580-general-e2-equivariant-steerable-cnns>
AUTHORS: Maurice Weiler, Gabriele Cesa
HIGHLIGHT: Here we give a general description of E(2)-equivariant convolutions in the framework of Steerable CNNs.
- 1286, TITLE: Characterization and Learning of Causal Graphs with Latent Variables from Soft Interventions
<https://papers.nips.cc/paper/9581-characterization-and-learning-of-causal-graphs-with-latent-variables-from-soft-interventions>
AUTHORS: Murat Kocaoglu, Amin Jaber, Karthikeyan Shanmugam, Elias Bareinboim
HIGHLIGHT: In this paper, we investigate the more general scenario where multiple observational and experimental distributions are available.
- 1287, TITLE: Structure Learning with Side Information: Sample Complexity
<https://papers.nips.cc/paper/9582-structure-learning-with-side-information-sample-complexity>
AUTHORS: Saurabh Sihag, Ali Tajer
HIGHLIGHT: This paper focuses on Ising graphical models, and considers the problem of simultaneously learning the structures of two $\{\text{sl partially}\}$ similar graphs, where any inference about the structure of one graph offers side information for the other graph.
- 1288, TITLE: Untangling in Invariant Speech Recognition
<https://papers.nips.cc/paper/9583-untangling-in-invariant-speech-recognition>
AUTHORS: Cory Stephenson, Jenelle Feather, Suchismita Padhy, Oguz Elibol, Hanlin Tang, Josh McDermott, SueYeon Chung
HIGHLIGHT: In this work, we employ a recently developed statistical mechanical theory that connects geometric properties of network representations and the separability of classes to probe how information is untangled within neural networks trained to recognize speech.
- 1289, TITLE: Flexible information routing in neural populations through stochastic comodulation
<https://papers.nips.cc/paper/9584-flexible-information-routing-in-neural-populations-through-stochastic-comodulation>
AUTHORS: Caroline Haimerl, Cristina Savin, Eero Simoncelli
HIGHLIGHT: Here, we propose a novel solution based on functionally-targeted stochastic modulation.
- 1290, TITLE: Generalization Bounds in the Predict-then-Optimize Framework
<https://papers.nips.cc/paper/9585-generalization-bounds-in-the-predict-then-optimize-framework>
AUTHORS: Othman El Balghiti, Adam Elmachtoub, Paul Grigas, Ambuj Tewari
HIGHLIGHT: In this work, we provide an assortment of generalization bounds for the SPO loss function.
- 1291, TITLE: Categorized Bandits
<https://papers.nips.cc/paper/9586-categorized-bandits>
AUTHORS: Matthieu Jedor, Vianney Perchet, Jonathan Louedec
HIGHLIGHT: We introduce a new stochastic multi-armed bandit setting where arms are grouped inside "ordered" categories.
- 1292, TITLE: Worst-Case Regret Bounds for Exploration via Randomized Value Functions
<https://papers.nips.cc/paper/9587-worst-case-regret-bounds-for-exploration-via-randomized-value-functions>
AUTHORS: Daniel Russo
HIGHLIGHT: By providing a worst-case regret bound for tabular finite-horizon Markov decision processes, we show that planning with respect to these randomized value functions can induce provably efficient exploration.
- 1293, TITLE: Efficient characterization of electrically evoked responses for neural interfaces
<https://papers.nips.cc/paper/9588-efficient-characterization-of-electrically-evoked-responses-for-neural-interfaces>
AUTHORS: Nishal Shah, Sasidhar Madugula, Pawel Hottowy, Alexander Sher, Alan Litke, Liam Paninski, E.J. Chichilnisky
HIGHLIGHT: This work tests the idea that using prior information from previous experiments and closed-loop measurements may greatly increase the efficiency of the neural interface.
- 1294, TITLE: Differentially Private Distributed Data Summarization under Covariate Shift
<https://papers.nips.cc/paper/9589-differentially-private-distributed-data-summarization-under-covariate-shift>
AUTHORS: Kanthi Sarpatwar, Karthikeyan Shanmugam, Venkata Sitaramagiridharganesh Ganapavarapu, Ashish Jagmohan, Roman Vaculin
HIGHLIGHT: We introduce a novel "noiseless" differentially private auctioning protocol, which may be of independent interest.

- 1295, TITLE: Hamiltonian descent for composite objectives
<https://papers.nips.cc/paper/9590-hamiltonian-descent-for-composite-objectives>
AUTHORS: Brendan O'Donoghue, Chris J. Maddison
HIGHLIGHT: In this paper we consider a convex optimization problem consisting of the sum of two convex functions, sometimes referred to as a composite objective, and we identify the duality gap to be the 'energy' of the system.
- 1296, TITLE: Implicit Regularization of Accelerated Methods in Hilbert Spaces
<https://papers.nips.cc/paper/9591-implicit-regularization-of-accelerated-methods-in-hilbert-spaces>
AUTHORS: Nicolò Pagliana, Lorenzo Rosasco
HIGHLIGHT: We study learning properties of accelerated gradient descent methods for linear least-squares in Hilbert spaces.
- 1297, TITLE: Non-Asymptotic Pure Exploration by Solving Games
<https://papers.nips.cc/paper/9592-non-asymptotic-pure-exploration-by-solving-games>
AUTHORS: Rémy Degenne, Wouter M. Koolen, Pierre Ménard
HIGHLIGHT: We interpret the optimisation problem as an unknown game, and propose sampling rules based on iterative strategies to estimate and converge to its saddle point.
- 1298, TITLE: Implicit Posterior Variational Inference for Deep Gaussian Processes
<https://papers.nips.cc/paper/9593-implicit-posterior-variational-inference-for-deep-gaussian-processes>
AUTHORS: Haibin YU, Yizhou Chen, Bryan Kian Hsiang Low, Patrick Jaillet, Zhongxiang Dai
HIGHLIGHT: This paper presents an implicit posterior variational inference (IPVI) framework for DGPs that can ideally recover an unbiased posterior belief and still preserve time efficiency.
- 1299, TITLE: Deep Multi-State Dynamic Recurrent Neural Networks Operating on Wavelet Based Neural Features for Robust Brain Machine Interfaces
<https://papers.nips.cc/paper/9594-deep-multi-state-dynamic-recurrent-neural-networks-operating-on-wavelet-based-neural-features-for-robust-brain-machine-interfaces>
AUTHORS: Benyamin Allahgholizadeh Haghi, Spencer Kellis, Sahil Shah, Maitreyi Ashok, Luke Bashford, Daniel Kramer, Brian Lee, Charles Liu, Richard Andersen, Azita Emami
HIGHLIGHT: We present a new deep multi-state Dynamic Recurrent Neural Network (DRNN) architecture for Brain Machine Interface (BMI) applications.
- 1300, TITLE: Censored Semi-Bandits: A Framework for Resource Allocation with Censored Feedback
<https://papers.nips.cc/paper/9595-censored-semi-bandits-a-framework-for-resource-allocation-with-censored-feedback>
AUTHORS: Arun Verma, Manjesh Hanawal, Arun Rajkumar, Raman Sankaran
HIGHLIGHT: In this paper, we study Censored Semi-Bandits, a novel variant of the semi-bandits problem.
- 1301, TITLE: Cormorant: Covariant Molecular Neural Networks
<https://papers.nips.cc/paper/9596-cormorant-covariant-molecular-neural-networks>
AUTHORS: Brandon Anderson, Truong Son Hy, Risi Kondor
HIGHLIGHT: We propose Cormorant, a rotationally covariant neural network architecture for learning the behavior and properties of complex many-body physical systems.
- 1302, TITLE: Reverse KL-Divergence Training of Prior Networks: Improved Uncertainty and Adversarial Robustness
<https://papers.nips.cc/paper/9597-reverse-kl-divergence-training-of-prior-networks-improved-uncertainty-and-adversarial-robustness>
AUTHORS: Andrey Malinin, Mark Gales
HIGHLIGHT: First, we show that the appropriate training criterion for Prior Networks is the reverse KL-divergence between Dirichlet distributions. This addresses issues in the nature of the training data target distributions, enabling prior networks to be successfully trained on classification tasks with arbitrarily many classes, as well as improving out-of-distribution detection performance. Second, taking advantage of this new training criterion, this paper investigates using Prior Networks to detect adversarial attacks and proposes a generalized form of adversarial training.
- 1303, TITLE: Reflection Separation using a Pair of Unpolarized and Polarized Images
<https://papers.nips.cc/paper/9598-reflection-separation-using-a-pair-of-unpolarized-and-polarized-images>
AUTHORS: Youwei Lyu, Zhaopeng Cui, Si Li, Marc Pollefeys, Boxin Shi
HIGHLIGHT: In this paper, we propose to exploit physical constraints from a pair of unpolarized and polarized images to separate reflection and transmission layers.
- 1304, TITLE: Policy Poisoning in Batch Reinforcement Learning and Control

<https://papers.nips.cc/paper/9599-policy-poisoning-in-batch-reinforcement-learning-and-control>

AUTHORS: Yuzhe Ma, Xuezhou Zhang, Wen Sun, Jerry Zhu
HIGHLIGHT: We present a unified framework for solving batch policy poisoning attacks, and instantiate the attack on two standard victims: tabular certainty equivalence learner in reinforcement learning and linear quadratic regulator in control.

1305, TITLE: Low-Complexity Nonparametric Bayesian Online Prediction with Universal Guarantees

<https://papers.nips.cc/paper/9600-low-complexity-nonparametric-bayesian-online-prediction-with-universal-guarantees>

AUTHORS: Alix LHERITIER, Frederic Cazals
HIGHLIGHT: We propose a novel nonparametric online predictor for discrete labels conditioned on multivariate continuous features.

1306, TITLE: Pure Exploration with Multiple Correct Answers

<https://papers.nips.cc/paper/9601-pure-exploration-with-multiple-correct-answers>

AUTHORS: Remy Degenne, Wouter M. Koolen
HIGHLIGHT: We present a new algorithm which extends Track-and-Stop to the multiple-answer case and has asymptotic sample complexity matching the lower bound.

1307, TITLE: Explaining Landscape Connectivity of Low-cost Solutions for Multilayer Nets

<https://papers.nips.cc/paper/9602-explaining-landscape-connectivity-of-low-cost-solutions-for-multilayer-nets>

AUTHORS: Rohith Kuditipudi, Xiang Wang, Holden Lee, Yi Zhang, Zhiyuan Li, Wei Hu, Rong Ge, Sanjeev Arora
HIGHLIGHT: We give mathematical explanations for this phenomenon, assuming generic properties (such as dropout stability and noise stability) of well-trained deep nets, which have previously been identified as part of understanding the generalization properties of deep nets.

1308, TITLE: On the Fairness of Disentangled Representations

<https://papers.nips.cc/paper/9603-on-the-fairness-of-disentangled-representations>

AUTHORS: Francesco Locatello, Gabriele Abbati, Thomas Rainforth, Stefan Bauer, Bernhard Schölkopf, Olivier Bachem
HIGHLIGHT: In this paper, we investigate the usefulness of different notions of disentanglement for improving the fairness of downstream prediction tasks based on representations.

1309, TITLE: Compiler Auto-Vectorization with Imitation Learning

<https://papers.nips.cc/paper/9604-compiler-auto-vectorization-with-imitation-learning>

AUTHORS: Charith Mendis, Cambridge Yang, Yewen Pu, Dr. Saman Amarasinghe, Michael Carbin
HIGHLIGHT: In this work, we explore whether it is feasible to imitate optimal decisions made by their ILP solution by fitting a graph neural network policy.

1310, TITLE: A Generalized Algorithm for Multi-Objective Reinforcement Learning and Policy Adaptation

<https://papers.nips.cc/paper/9605-a-generalized-algorithm-for-multi-objective-reinforcement-learning-and-policy-adaptation>

AUTHORS: Runzhe Yang, Xingyuan Sun, Karthik Narasimhan
HIGHLIGHT: We introduce a new algorithm for multi-objective reinforcement learning (MORL) with linear preferences, with the goal of enabling few-shot adaptation to new tasks.

1311, TITLE: Exact Gaussian Processes on a Million Data Points

<https://papers.nips.cc/paper/9606-exact-gaussian-processes-on-a-million-data-points>

AUTHORS: Ke Wang, Geoff Pleiss, Jacob Gardner, Stephen Tyree, Kilian Q. Weinberger, Andrew Gordon Wilson
HIGHLIGHT: In this paper, we develop a scalable approach for exact GPs that leverages multi-GPU parallelization and methods like linear conjugate gradients, accessing the kernel matrix only through matrix multiplication.

1312, TITLE: Bayesian Layers: A Module for Neural Network Uncertainty

<https://papers.nips.cc/paper/9607-bayesian-layers-a-module-for-neural-network-uncertainty>

AUTHORS: Dustin Tran, Mike Dusenberry, Mark van der Wilk, Danijar Hafner
HIGHLIGHT: We describe Bayesian Layers, a module designed for fast experimentation with neural network uncertainty.

1313, TITLE: Learning Compositional Neural Programs with Recursive Tree Search and Planning

<https://papers.nips.cc/paper/9608-learning-compositional-neural-programs-with-recursive-tree-search-and-planning>

AUTHORS: Thomas PIERROT, Guillaume Ligner, Scott E. Reed, Olivier Sigaud, Nicolas Perrin, Alexandre Laterre, David Kas, Karim Beguir, Nando de Freitas
HIGHLIGHT: We propose a novel reinforcement learning algorithm, AlphaNPI, that incorporates the strengths of Neural Programmer-Interpreters (NPI) and AlphaZero.

- 1314, TITLE: Nonparametric Contextual Bandits in Metric Spaces with Unknown Metric
<https://papers.nips.cc/paper/9609-nonparametric-contextual-bandits-in-metric-spaces-with-unknown-metric>
AUTHORS: Nirandika Wanigasekara, Christina Yu
HIGHLIGHT: We present a novel algorithm which learns data-driven similarities amongst the arms, in order to implement adaptive partitioning of the context-arm space for more efficient learning.
- 1315, TITLE: Qsparse-local-SGD: Distributed SGD with Quantization, Sparsification and Local Computations
<https://papers.nips.cc/paper/9610-qsparse-local-sgd-distributed-sgd-with-quantization-sparsification-and-local-computations>
AUTHORS: Debraj Basu, Deepesh Data, Can Karakus, Suhas Diggavi
HIGHLIGHT: In this paper we propose Qsparse-local-SGD algorithm, which combines aggressive sparsification with quantization and local computation along with error compensation, by keeping track of the difference between the true and compressed gradients.
- 1316, TITLE: Likelihood Ratios for Out-of-Distribution Detection
<https://papers.nips.cc/paper/9611-likelihood-ratios-for-out-of-distribution-detection>
AUTHORS: Jie Ren, Peter J. Liu, Emily Fertig, Jasper Snoek, Ryan Poplin, Mark Depristo, Joshua Dillon, Balaji Lakshminarayanan
HIGHLIGHT: We propose a likelihood ratio method for deep generative models which effectively corrects for these confounding background statistics.
- 1317, TITLE: Discrete Flows: Invertible Generative Models of Discrete Data
<https://papers.nips.cc/paper/9612-discrete-flows-invertible-generative-models-of-discrete-data>
AUTHORS: Dustin Tran, Keyon Vafa, Kumar Agrawal, Laurent Dinh, Ben Poole
HIGHLIGHT: In this paper, we show that flows can in fact be extended to discrete events---and under a simple change-of-variables formula not requiring log-determinant-Jacobian computations.
- 1318, TITLE: A Self Validation Network for Object-Level Human Attention Estimation
<https://papers.nips.cc/paper/9613-a-self-validation-network-for-object-level-human-attention-estimation>
AUTHORS: Zehua Zhang, Chen Yu, David Crandall
HIGHLIGHT: In this paper, we propose a novel unified model that incorporates both spatial and temporal evidence in identifying as well as locating the attended object in firstperson videos.
- 1319, TITLE: Model Selection for Contextual Bandits
<https://papers.nips.cc/paper/9614-model-selection-for-contextual-bandits>
AUTHORS: Dylan J. Foster, Akshay Krishnamurthy, Haipeng Luo
HIGHLIGHT: We introduce the problem of model selection for contextual bandits, where a learner must adapt to the complexity of the optimal policy while balancing exploration and exploitation.
- 1320, TITLE: Sliced Gromov-Wasserstein
<https://papers.nips.cc/paper/9615-sliced-gromov-wasserstein>
AUTHORS: Vayer Titouan, R?mi Flamary, Nicolas Courty, Romain Tavenard, Laetitia Chapel
HIGHLIGHT: This paper proposes a new divergence based on GW akin to SW.
- 1321, TITLE: Towards Practical Alternating Least-Squares for CCA
<https://papers.nips.cc/paper/9616-towards-practical-alternating-least-squares-for-cca>
AUTHORS: Zhiqiang Xu, Ping Li
HIGHLIGHT: To promote the practical use of ALS for CCA, we propose truly alternating least-squares.
- 1322, TITLE: Deep Leakage from Gradients
<https://papers.nips.cc/paper/9617-deep-leakage-from-gradients>
AUTHORS: Ligeng Zhu, Zhijian Liu, Song Han
HIGHLIGHT: However, in this paper, we show that we can obtain the private training set from the publicly shared gradients.
- 1323, TITLE: Invariance-inducing regularization using worst-case transformations suffices to boost accuracy and spatial robustness
<https://papers.nips.cc/paper/9618-invariance-inducing-regularization-using-worst-case-transformations-suffices-to-boost-accuracy-and-spatial-robustness>
AUTHORS: Fanny Yang, Zuowen Wang, Christina Heinze-Deml
HIGHLIGHT: This work provides theoretical and empirical evidence that invariance-inducing regularizers can increase predictive accuracy for worst-case spatial transformations (spatial robustness).

1324, TITLE: Algorithm-Dependent Generalization Bounds for Overparameterized Deep Residual Networks
<https://papers.nips.cc/paper/9619-algorithm-dependent-generalization-bounds-for-overparameterized-deep-residual-networks>
AUTHORS: Spencer Frei, Yuan Cao, Quanquan Gu
HIGHLIGHT: In this work, we analyze overparameterized deep residual networks trained by gradient descent following random initialization, and demonstrate that (i) the class of networks learned by gradient descent constitutes a small subset of the entire neural network function class, and (ii) this subclass of networks is sufficiently large to guarantee small training error.

1325, TITLE: Value Function in Frequency Domain and the Characteristic Value Iteration Algorithm
<https://papers.nips.cc/paper/9620-value-function-in-frequency-domain-and-the-characteristic-value-iteration-algorithm>
AUTHORS: Amir-massoud Farahmand
HIGHLIGHT: It presents a new representational framework to maintain the uncertainty of returns and provides mathematical tools to compute it.

1326, TITLE: Icebreaker: Element-wise Efficient Information Acquisition with a Bayesian Deep Latent Gaussian Model
<https://papers.nips.cc/paper/9621-icebreaker-element-wise-efficient-information-acquisition-with-a-bayesian-deep-latent-gaussian-model>
AUTHORS: Wenbo Gong, Sebastian Tschiatschek, Sebastian Nowozin, Richard E. Turner, Jos? Miguel Hern?andez-Lobato, Cheng Zhang
HIGHLIGHT: In this paper, we address the ice-start problem, i.e., the challenge of deploying machine learning models when only a little or no training data is initially available, and acquiring each feature element of data is associated with costs.

1327, TITLE: Algorithmic Guarantees for Inverse Imaging with Untrained Network Priors
<https://papers.nips.cc/paper/9622-algorithmic-guarantees-for-inverse-imaging-with-untrained-network-priors>
AUTHORS: Gauri Jagatap, Chinmay Hegde
HIGHLIGHT: Specifically, we consider the problem of solving linear inverse problems, such as compressive sensing, as well as non-linear problems, such as compressive phase retrieval.

1328, TITLE: Planning with Goal-Conditioned Policies
<https://papers.nips.cc/paper/9623-planning-with-goal-conditioned-policies>
AUTHORS: Soroush Nasiriany, Vitchyr Pong, Steven Lin, Sergey Levine
HIGHLIGHT: We show that goal-conditioned policies learned with RL can be incorporated into planning, such that a planner can focus on which states to reach, rather than how those states are reached.

1329, TITLE: Don't take it lightly: Phasing optical random projections with unknown operators
<https://papers.nips.cc/paper/9624-dont-take-it-lightly-phasing-optical-random-projections-with-unknown-operators>
AUTHORS: Sidharth Gupta, Remi Gribonval, Laurent Daudet, Ivan Dokmanic
HIGHLIGHT: In this paper we tackle the problem of recovering the phase of complex linear measurements when only magnitude information is available and we control the input.

1330, TITLE: Generating Diverse High-Fidelity Images with VQ-VAE-2
<https://papers.nips.cc/paper/9625-generating-diverse-high-fidelity-images-with-vq-vae-2>
AUTHORS: Ali Razavi, Aaron van den Oord, Oriol Vinyals
HIGHLIGHT: We explore the use of Vector Quantized Variational AutoEncoder (VQ-VAE) models for large scale image generation.

1331, TITLE: Generalized Matrix Means for Semi-Supervised Learning with Multilayer Graphs
<https://papers.nips.cc/paper/9626-generalized-matrix-means-for-semi-supervised-learning-with-multilayer-graphs>
AUTHORS: Pedro Mercado, Francesco Tudisco, Matthias Hein
HIGHLIGHT: We propose a regularizer based on the generalized matrix mean, which is a one-parameter family of matrix means that includes the arithmetic, geometric and harmonic means as particular cases.

1332, TITLE: Online Optimal Control with Linear Dynamics and Predictions: Algorithms and Regret Analysis
<https://papers.nips.cc/paper/9627-online-optimal-control-with-linear-dynamics-and-predictions-algorithms-and-regret-analysis>
AUTHORS: Yingying Li, Xin Chen, Na Li
HIGHLIGHT: We design online algorithms, Receding Horizon Gradient-based Control (RHGC), that utilize the predictions through finite steps of gradient computations.

1333, TITLE: Missing Not at Random in Matrix Completion: The Effectiveness of Estimating Missingness Probabilities Under a Low Nuclear Norm Assumption

- <https://papers.nips.cc/paper/9628-missing-not-at-random-in-matrix-completion-the-effectiveness-of-estimating-missingness-probabilities-under-a-low-nuclear-norm-assumption>
AUTHORS: Wei Ma, George H. Chen
HIGHLIGHT: In this paper, we suggest a simple approach to estimating these probabilities that avoids these shortcomings.
- 1334, TITLE: MelGAN: Generative Adversarial Networks for Conditional Waveform Synthesis
<https://papers.nips.cc/paper/9629-melgan-generative-adversarial-networks-for-conditional-waveform-synthesis>
AUTHORS: Kundan Kumar, Rithesh Kumar, Thibault de Boissiere, Lucas Gestin, Wei Zhen Teoh, Jose Sotelo, Alexandre de Br?bisson, Yoshua Bengio, Aaron C. Courville
HIGHLIGHT: In this paper, we show that it is possible to train GANs reliably to generate high quality coherent waveforms by introducing a set of architectural changes and simple training techniques.
- 1335, TITLE: Offline Contextual Bandits with High Probability Fairness Guarantees
<https://papers.nips.cc/paper/9630-offline-contextual-bandits-with-high-probability-fairness-guarantees>
AUTHORS: Blossom Metevier, Stephen Giguere, Sarah Brockman, Ari Kobren, Yuriy Brun, Emma Brunskill, Philip S. Thomas
HIGHLIGHT: We present RobinHood, an offline contextual bandit algorithm designed to satisfy a broad family of fairness constraints.
- 1336, TITLE: Solving a Class of Non-Convex Min-Max Games Using Iterative First Order Methods
<https://papers.nips.cc/paper/9631-solving-a-class-of-non-convex-min-max-games-using-iterative-first-order-methods>
AUTHORS: Maher Nouiehed, Maziar Sanjabi, Tianjian Huang, Jason D. Lee, Meisam Razaviyayn
HIGHLIGHT: In this paper, we study the problem in the non-convex regime and show that an ϵ -first order stationary point of the game can be computed when one of the player's objective can be optimized to global optimality efficiently.
- 1337, TITLE: Semantic-Guided Multi-Attention Localization for Zero-Shot Learning
<https://papers.nips.cc/paper/9632-semantic-guided-multi-attention-localization-for-zero-shot-learning>
AUTHORS: Yizhe Zhu, Jianwen Xie, Zhiqiang Tang, Xi Peng, Ahmed Elgammal
HIGHLIGHT: In this paper, we study the significance of the discriminative region localization.
- 1338, TITLE: Interpreting and improving natural-language processing (in machines) with natural language-processing (in the brain)
<https://papers.nips.cc/paper/9633-interpreting-and-improving-natural-language-processing-in-machines-with-natural-language-processing-in-the-brain>
AUTHORS: Mariya Toneva, Leila Wehbe
HIGHLIGHT: We propose here a novel interpretation approach that relies on the only processing system we have that does understand language: the human brain.
- 1339, TITLE: Function-Space Distributions over Kernels
<https://papers.nips.cc/paper/9634-function-space-distributions-over-kernels>
AUTHORS: Gregory Benton, Wesley J. Maddox, Jayson Salkey, Julio Albinati, Andrew Gordon Wilson
HIGHLIGHT: In this paper, we develop functional kernel learning (FKL) to directly infer functional posteriors over kernels.
- 1340, TITLE: The Step Decay Schedule: A Near Optimal, Geometrically Decaying Learning Rate Procedure For Least Squares
<https://papers.nips.cc/paper/9635-the-step-decay-schedule-a-near-optimal-geometrically-decaying-learning-rate-procedure-for-least-squares>
AUTHORS: Rong Ge, Sham M. Kakade, Rahul Kidambi, Praneeth Netrapalli
HIGHLIGHT: Motivated by this observation, this work provides a detailed study of the following question: what rate is achievable using the final iterate of SGD for the streaming least squares regression problem with and without strong convexity?
- 1341, TITLE: Compositional Plan Vectors
<https://papers.nips.cc/paper/9636-compositional-plan-vectors>
AUTHORS: Coline Devin, Daniel Geng, Pieter Abbeel, Trevor Darrell, Sergey Levine
HIGHLIGHT: We introduce compositional plan vectors (CPVs) to enable a policy to perform compositions of tasks without additional supervision.
- 1342, TITLE: Locally Private Learning without Interaction Requires Separation
<https://papers.nips.cc/paper/9637-locally-private-learning-without-interaction-requires-separation>
AUTHORS: Amit Daniely, Vitaly Feldman
HIGHLIGHT: We consider learning under the constraint of local differential privacy (LDP).

- 1343, TITLE: Robust Bi-Tempered Logistic Loss Based on Bregman Divergences
<https://papers.nips.cc/paper/9638-robust-bi-tempered-logistic-loss-based-on-bregman-divergences>
AUTHORS: Ehsan Amid, Manfred K. K. Warmuth, Rohan Anil, Tomer Koren
HIGHLIGHT: We introduce a temperature into the exponential function and replace the softmax output layer of the neural networks by a high-temperature generalization.
- 1344, TITLE: Computational Separations between Sampling and Optimization
<https://papers.nips.cc/paper/9639-computational-separations-between-sampling-and-optimization>
AUTHORS: Kunal Talwar
HIGHLIGHT: We present a simpler and stronger separation.
- 1345, TITLE: Surfing: Iterative Optimization Over Incrementally Trained Deep Networks
<https://papers.nips.cc/paper/9640-surfing-iterative-optimization-over-incrementally-trained-deep-networks>
AUTHORS: Ganlin Song, Zhou Fan, John Lafferty
HIGHLIGHT: We investigate a sequential optimization procedure to minimize the empirical risk functional $f_{\hat{\theta}}(x) = \frac{1}{2} \|G_{\hat{\theta}}(x) - y\|^2$ for certain families of deep networks $G_{\theta}(x)$.
- 1346, TITLE: Learning to Optimize in Swarms
<https://papers.nips.cc/paper/9641-learning-to-optimize-in-swarms>
AUTHORS: Yue Cao, Tianlong Chen, Zhangyang Wang, Yang Shen
HIGHLIGHT: To overcome the limitations, we propose a meta-optimizer that learns in the algorithmic space of both point-based and population-based optimization algorithms.
- 1347, TITLE: On Human-Aligned Risk Minimization
<https://papers.nips.cc/paper/9642-on-human-aligned-risk-minimization>
AUTHORS: Liu Leqi, Adarsh Prasad, Pradeep K. Ravikumar
HIGHLIGHT: In this paper, we pose the following simple question: in contrast to minimizing expected loss, could we minimize a better human-aligned risk measure?
- 1348, TITLE: Semi-Parametric Efficient Policy Learning with Continuous Actions
<https://papers.nips.cc/paper/9643-semi-parametric-efficient-policy-learning-with-continuous-actions>
AUTHORS: Victor Chernozhukov, Mert Demirer, Greg Lewis, Vasilis Syrgkanis
HIGHLIGHT: We propose a doubly robust off-policy estimate for this setting and show that off-policy optimization based on this doubly robust estimate is robust to estimation errors of the policy function or the regression model.
- 1349, TITLE: Multi-task Learning for Aggregated Data using Gaussian Processes
<https://papers.nips.cc/paper/9644-multi-task-learning-for-aggregated-data-using-gaussian-processes>
AUTHORS: Fariba Yousefi, Michael T. Smith, Mauricio Alvarez
HIGHLIGHT: In this paper, we present a novel multi-task learning model based on Gaussian processes for joint learning of variables that have been aggregated at different input scales.
- 1350, TITLE: Minimal Variance Sampling in Stochastic Gradient Boosting
<https://papers.nips.cc/paper/9645-minimal-variance-sampling-in-stochastic-gradient-boosting>
AUTHORS: Bulat Ibragimov, Gleb Gusev
HIGHLIGHT: In this paper, we formulate the problem of randomization in SGB in terms of optimization of sampling probabilities to maximize the estimation accuracy of split scoring used to train decision trees.
- 1351, TITLE: Beyond the Single Neuron Convex Barrier for Neural Network Certification
<https://papers.nips.cc/paper/9646-beyond-the-single-neuron-convex-barrier-for-neural-network-certification>
AUTHORS: Gagandeep Singh, Rupanshu Ganvir, Markus P. Schel, Martin Vechev
HIGHLIGHT: We propose a new parametric framework, called k-ReLU, for computing precise and scalable convex relaxations used to certify neural networks.
- 1352, TITLE: An Algorithm to Learn Polytree Networks with Hidden Nodes
<https://papers.nips.cc/paper/9647-an-algorithm-to-learn-polytree-networks-with-hidden-nodes>
AUTHORS: Firoozeh Sepehr, Donatello Materassi
HIGHLIGHT: In this article, we develop an algorithm to exactly recover graphical models of random variables with underlying polytree structures when the latent nodes satisfy specific degree conditions.

- 1353, TITLE: Efficiently Learning Fourier Sparse Set Functions
<https://papers.nips.cc/paper/9648-efficiently-learning-fourier-sparse-set-functions>
AUTHORS: Andisheh Amrollahi, Amir Zandieh, Michael Kapralov, Andreas Krause
HIGHLIGHT: In this paper we consider the problem of efficiently learning set functions that are defined over a ground set of size n and that are sparse (say k -sparse) in the Fourier domain.
- 1354, TITLE: Projected Stein Variational Newton: A Fast and Scalable Bayesian Inference Method in High Dimensions
<https://papers.nips.cc/paper/9649-projected-stein-variational-newton-a-fast-and-scalable-bayesian-inference-method-in-high-dimensions>
AUTHORS: Peng Chen, Keyi Wu, Joshua Chen, Tom O'Leary-Roseberry, Omar Ghattas
HIGHLIGHT: We propose a projected Stein variational Newton (pSVN) method for high-dimensional Bayesian inference.
- 1355, TITLE: Invariance and identifiability issues for word embeddings
<https://papers.nips.cc/paper/9650-invariance-and-identifiability-issues-for-word-embeddings>
AUTHORS: Rachel Carrington, Karthik Bharath, Simon Preston
HIGHLIGHT: We provide a formal treatment of the above identifiability issue, present some numerical examples, and discuss possible resolutions.
- 1356, TITLE: Generalization Error Analysis of Quantized Compressive Learning
<https://papers.nips.cc/paper/9651-generalization-error-analysis-of-quantized-compressive-learning>
AUTHORS: Xiaoyun Li, Ping Li
HIGHLIGHT: In this paper, we consider the learning problem where the projected data is further compressed by scalar quantization, which is called quantized compressive learning.
- 1357, TITLE: Multi-Criteria Dimensionality Reduction with Applications to Fairness
<https://papers.nips.cc/paper/9652-multi-criteria-dimensionality-reduction-with-applications-to-fairness>
AUTHORS: Uthaiapon Tantipongpipat, Samira Samadi, Mohit Singh, Jamie H. Morgenstern, Santosh Vempala
HIGHLIGHT: In this paper, we introduce the multi-criteria dimensionality reduction problem where we are given multiple objectives that need to be optimized simultaneously.
- 1358, TITLE: Efficient Rematerialization for Deep Networks
<https://papers.nips.cc/paper/9653-efficient-rematerialization-for-deep-networks>
AUTHORS: Ravi Kumar, Manish Purohit, Zoya Svitkina, Erik Vee, Joshua Wang
HIGHLIGHT: In this work we consider the rematerialization problem and devise efficient algorithms that use structural characterizations of computation graphs---treewidth and pathwidth---to obtain provably efficient rematerialization schedules.
- 1359, TITLE: Mo' States Mo' Problems: Emergency Stop Mechanisms from Observation
<https://papers.nips.cc/paper/9654-mo-states-mo-problems-emergency-stop-mechanisms-from-observation>
AUTHORS: Samuel Ainsworth, Matt Barnes, Siddhartha Srinivasa
HIGHLIGHT: We develop a simple technique using emergency stops (e-stops) to exploit this phenomenon.
- 1360, TITLE: Machine Learning Estimation of Heterogeneous Treatment Effects with Instruments
<https://papers.nips.cc/paper/9655-machine-learning-estimation-of-heterogeneous-treatment-effects-with-instruments>
AUTHORS: Vasilis Syrgkanis, Victor Lei, Miruna Oprescu, Maggie Hei, Keith Battocchi, Greg Lewis
HIGHLIGHT: We develop a statistical learning approach to the estimation of heterogeneous effects, reducing the problem to the minimization of an appropriate loss function that depends on a set of auxiliary models (each corresponding to a separate prediction task).
- 1361, TITLE: Understanding Sparse JL for Feature Hashing
<https://papers.nips.cc/paper/9656-understanding-sparse-jl-for-feature-hashing>
AUTHORS: Meena Jagadeesan
HIGHLIGHT: In this paper, we demonstrate the benefits of using sparsity s greater than 1 in sparse JL on feature vectors.
- 1362, TITLE: Text-Based Interactive Recommendation via Constraint-Augmented Reinforcement Learning
<https://papers.nips.cc/paper/9657-text-based-interactive-recommendation-via-constraint-augmented-reinforcement-learning>
AUTHORS: Ruiyi Zhang, Tong Yu, Yilin Shen, Hongxia Jin, Changyou Chen
HIGHLIGHT: To alleviate this issue, we propose a novel constraint-augmented reinforcement learning (RL) framework to efficiently incorporate user preferences over time.

- 1363, TITLE: Flexible Modeling of Diversity with Strongly Log-Concave Distributions
<https://papers.nips.cc/paper/9658-flexible-modeling-of-diversity-with-strongly-log-concave-distributions>
AUTHORS: Joshua Robinson, Suvrit Sra, Stefanie Jegelka
HIGHLIGHT: We propose SLC as the right extension of SR that enables easier, more intuitive control over diversity, illustrating this via examples of practical importance.
- 1364, TITLE: Momentum-Based Variance Reduction in Non-Convex SGD
<https://papers.nips.cc/paper/9659-momentum-based-variance-reduction-in-non-convex-sgd>
AUTHORS: Ashok Cutkosky, Francesco Orabona
HIGHLIGHT: We present a new algorithm, STORM, that does not require any batches and makes use of adaptive learning rates, enabling simpler implementation and less hyperparameter tuning.
- 1365, TITLE: Search on the Replay Buffer: Bridging Planning and Reinforcement Learning
<https://papers.nips.cc/paper/9660-search-on-the-replay-buffer-bridging-planning-and-reinforcement-learning>
AUTHORS: Ben Eysenbach, Russ R. Salakhutdinov, Sergey Levine
HIGHLIGHT: We introduce a general-purpose control algorithm that combines the strengths of planning and reinforcement learning to effectively solve these tasks.
- 1366, TITLE: Can Unconditional Language Models Recover Arbitrary Sentences?
<https://papers.nips.cc/paper/9661-can-unconditional-language-models-recover-arbitrary-sentences>
AUTHORS: Nishant Subramani, Samuel Bowman, Kyunghyun Cho
HIGHLIGHT: To do this, we introduce a pair of effective, complementary methods for feeding representations into pretrained unconditional language models and a corresponding set of methods to map sentences into and out of this representation space, the reparametrized sentence space.
- 1367, TITLE: Group Retention when Using Machine Learning in Sequential Decision Making: the Interplay between User Dynamics and Fairness
<https://papers.nips.cc/paper/9662-group-retention-when-using-machine-learning-in-sequential-decision-making-the-interplay-between-user-dynamics-and-fairness>
AUTHORS: Xueru Zhang, Mohammadmahdi Khaliligarekani, Cem Tekin, mingyan liu
HIGHLIGHT: In this study, we seek to understand the interplay between ML decisions and the underlying group representation, how they evolve in a sequential framework, and how the use of fairness criteria plays a role in this process.
- 1368, TITLE: Faster width-dependent algorithm for mixed packing and covering LPs
<https://papers.nips.cc/paper/9663-faster-width-dependent-algorithm-for-mixed-packing-and-covering-lps>
AUTHORS: Digvijay Boob, Saurabh Sawlani, Di Wang
HIGHLIGHT: In this paper, we give a faster width-dependent algorithm for mixed packing-covering LPs.
- 1369, TITLE: Flattening a Hierarchical Clustering through Active Learning
<https://papers.nips.cc/paper/9664-flattening-a-hierarchical-clustering-through-active-learning>
AUTHORS: Fabio Vitale, Anand Rajagopalan, Claudio Gentile
HIGHLIGHT: We investigate active learning by pairwise similarity over the leaves of trees originating from hierarchical clustering procedures.
- 1370, TITLE: DeepWave: A Recurrent Neural-Network for Real-Time Acoustic Imaging
<https://papers.nips.cc/paper/9665-deepwave-a-recurrent-neural-network-for-real-time-acoustic-imaging>
AUTHORS: Matthieu SIMEONI, Sepand Kashani, Paul Hurley, Martin Vetterli
HIGHLIGHT: We propose a recurrent neural-network for real-time reconstruction of acoustic camera spherical maps.
- 1371, TITLE: Certifying Geometric Robustness of Neural Networks
<https://papers.nips.cc/paper/9666-certifying-geometric-robustness-of-neural-networks>
AUTHORS: Mislav Balunovic, Maximilian Baader, Gagandeep Singh, Timon Gehr, Martin Vechev
HIGHLIGHT: In this work, we propose a new method to compute sound and asymptotically optimal linear relaxations for any composition of transformations.
- 1372, TITLE: Goal-conditioned Imitation Learning
<https://papers.nips.cc/paper/9667-goal-conditioned-imitation-learning>
AUTHORS: Yiming Ding, Carlos Florensa, Pieter Abbeel, Mariano Phielipp
HIGHLIGHT: In this work we propose a novel algorithm goalGAIL, which incorporates demonstrations to drastically speed up the convergence to a policy able to reach any goal, surpassing the performance of an agent trained with other Imitation Learning algorithms.

- 1373, TITLE: Robust exploration in linear quadratic reinforcement learning
<https://papers.nips.cc/paper/9668-robust-exploration-in-linear-quadratic-reinforcement-learning>
AUTHORS: Jack Umenberger, Mina Ferizbegovic, Thomas B. Sch?n, H?kan Hjalmarsson
HIGHLIGHT: We present a method, based on convex optimization, that accomplishes this task ‘robustly’, i.e., the worst-case cost, accounting for system uncertainty given the observed data, is minimized.
- 1374, TITLE: DRUM: End-To-End Differentiable Rule Mining On Knowledge Graphs
<https://papers.nips.cc/paper/9669-drum-end-to-end-differentiable-rule-mining-on-knowledge-graphs>
AUTHORS: Ali Sadeghian, Mohammadreza Armandpour, Patrick Ding, Daisy Zhe Wang
HIGHLIGHT: In this paper, we study the problem of learning probabilistic logical rules for inductive and interpretable link prediction.
- 1375, TITLE: Kernel Truncated Randomized Ridge Regression: Optimal Rates and Low Noise Acceleration
<https://papers.nips.cc/paper/9670-kernel-truncated-randomized-ridge-regression-optimal-rates-and-low-noise-acceleration>
AUTHORS: Kwang-Sung Jun, Ashok Cutkosky, Francesco Orabona
HIGHLIGHT: In this paper we consider the nonparametric least square regression in a Reproducing Kernel Hilbert Space (RKHS).
- 1376, TITLE: Input-Output Equivalence of Unitary and Contractive RNNs
<https://papers.nips.cc/paper/9671-input-output-equivalence-of-unitary-and-contractive-rnns>
AUTHORS: Melikasadat Emami, Mojtaba Sahraee Ardakan, Sundeep Rangan, Alyson K. Fletcher
HIGHLIGHT: This works shows that for any contractive RNN with ReLU activations, there is a URNN with at most twice the number of hidden states and the identical input-output mapping.
- 1377, TITLE: Hamiltonian Neural Networks
<https://papers.nips.cc/paper/9672-hamiltonian-neural-networks>
AUTHORS: Samuel Greydanus, Misko Dzamba, Jason Yosinski
HIGHLIGHT: In this paper, we draw inspiration from Hamiltonian mechanics to train models that learn and respect exact conservation laws in an unsupervised manner.
- 1378, TITLE: Preventing Gradient Attenuation in Lipschitz Constrained Convolutional Networks
<https://papers.nips.cc/paper/9673-preventing-gradient-attenuation-in-lipschitz-constrained-convolutional-networks>
AUTHORS: Qiyang Li, Saminul Haque, Cem Anil, James Lucas, Roger B. Grosse, Joern-Henrik Jacobsen
HIGHLIGHT: In particular, we present the Block Convolution Orthogonal Parameterization (BCOP), an expressive parameterization of orthogonal convolution operations.
- 1379, TITLE: Structured and Deep Similarity Matching via Structured and Deep Hebbian Networks
<https://papers.nips.cc/paper/9674-structured-and-deep-similarity-matching-via-structured-and-deep-hebbian-networks>
AUTHORS: Dina Obeid, Hugo Ramambason, Cengiz Pehlevan
HIGHLIGHT: In this paper, we introduce structured and deep similarity matching cost functions, and show how they can be optimized in a gradient-based manner by neural networks with local learning rules.
- 1380, TITLE: Understanding the Representation Power of Graph Neural Networks in Learning Graph Topology
<https://papers.nips.cc/paper/9675-understanding-the-representation-power-of-graph-neural-networks-in-learning-graph-topology>
AUTHORS: Nima Dehmamy, Albert-Laszlo Barabasi, Rose Yu
HIGHLIGHT: To deepen our understanding of graph neural networks, we investigate the representation power of Graph Convolutional Networks (GCN) through the looking glass of graph moments, a key property of graph topology encoding path of various lengths.
- 1381, TITLE: Multiple Futures Prediction
<https://papers.nips.cc/paper/9676-multiple-futures-prediction>
AUTHORS: Charlie Tang, Russ R. Salakhutdinov
HIGHLIGHT: Towards these goals, we introduce a probabilistic framework that efficiently learns latent variables to jointly model the multi-step future motions of agents in a scene.
- 1382, TITLE: Explicitly disentangling image content from translation and rotation with spatial-VAE
<https://papers.nips.cc/paper/9677-explicitly-disentangling-image-content-from-translation-and-rotation-with-spatial-vae>
AUTHORS: Tristan Bepler, Ellen Zhong, Kotaro Kelley, Edward Brignole, Bonnie Berger

HIGHLIGHT: We propose a method for explicitly disentangling image rotation and translation from other unstructured latent factors in a variational autoencoder (VAE) framework.

1383, TITLE: Power analysis of knockoff filters for correlated designs
<https://papers.nips.cc/paper/9678-power-analysis-of-knockoff-filters-for-correlated-designs>

AUTHORS: Jingbo Liu, Philippe Rigollet

HIGHLIGHT: In this work we study the case where the predictors have a general covariance matrix Σ .

1384, TITLE: A Kernel Loss for Solving the Bellman Equation
<https://papers.nips.cc/paper/9679-a-kernel-loss-for-solving-the-bellman-equation>

AUTHORS: Yihao Feng, Lihong Li, Qiang Liu

HIGHLIGHT: In this paper, we propose a novel loss function, which can be optimized using standard gradient-based methods with guaranteed convergence.

1385, TITLE: Low-Rank Bandit Methods for High-Dimensional Dynamic Pricing
<https://papers.nips.cc/paper/9680-low-rank-bandit-methods-for-high-dimensional-dynamic-pricing>

AUTHORS: Jonas W. Mueller, Vasilis Syrgkanis, Matt Taddy

HIGHLIGHT: We consider dynamic pricing with many products under an evolving but low-dimensional demand model.

1386, TITLE: Differential Privacy Has Disparate Impact on Model Accuracy
<https://papers.nips.cc/paper/9681-differential-privacy-has-disparate-impact-on-model-accuracy>

AUTHORS: Eugene Bagdasaryan, Omid Poursaeed, Vitaly Shmatikov

HIGHLIGHT: We demonstrate that in the neural networks trained using differentially private stochastic gradient descent (DP-SGD), this cost is not borne equally: accuracy of DP models drops much more for the underrepresented classes and subgroups.

1387, TITLE: Riemannian batch normalization for SPD neural networks
<https://papers.nips.cc/paper/9682-riemannian-batch-normalization-for-spd-neural-networks>

AUTHORS: Daniel Brooks, Olivier Schwander, Frederic Barbaresco, Jean-Yves Schneider, Matthieu Cord

HIGHLIGHT: In our article, we introduce a Riemannian batch normalization (batch-norm) algorithm, which generalizes the one used in Euclidean nets.

1388, TITLE: Neural Taskonomy: Inferring the Similarity of Task-Derived Representations from Brain Activity
<https://papers.nips.cc/paper/9683-neural-taskonomy-inferring-the-similarity-of-task-derived-representations-from-brain-activity>

AUTHORS: Yuan Wang, Michael Tarr, Leila Wehbe

HIGHLIGHT: To address this problem, we used learned representations drawn from 21 computer vision tasks to construct encoding models for predicting brain responses from BOLD5000—a large-scale dataset comprised of fMRI scans collected while observers viewed over 5000 naturalistic scene and object images.

1389, TITLE: Stacked Capsule Autoencoders
<https://papers.nips.cc/paper/9684-stacked-capsule-autoencoders>

AUTHORS: Adam Kosiorek, Sara Sabour, Yee Whye Teh, Geoffrey E. Hinton

HIGHLIGHT: We introduce an unsupervised capsule autoencoder (SCAE), which explicitly uses geometric relationships between parts to reason about objects.

1390, TITLE: Learning Reward Machines for Partially Observable Reinforcement Learning
<https://papers.nips.cc/paper/9685-learning-reward-machines-for-partially-observable-reinforcement-learning>

AUTHORS: Rodrigo Toro Icarte, Ethan Waldie, Toryn Klassen, Rick Valenzano, Margarita Castro, Sheila McIlraith

HIGHLIGHT: We pose the task of learning RMs as a discrete optimization problem where the objective is to find an RM that decomposes the problem into a set of subproblems such that the combination of their optimal memoryless policies is an optimal policy for the original problem.

1391, TITLE: Learning Representations by Maximizing Mutual Information Across Views
<https://papers.nips.cc/paper/9686-learning-representations-by-maximizing-mutual-information-across-views>

AUTHORS: Philip Bachman, R Devon Hjelm, William Buchwalter

HIGHLIGHT: We propose an approach to self-supervised representation learning based on maximizing mutual information between features extracted from multiple views of a shared context.

1392, TITLE: Amortized Bethe Free Energy Minimization for Learning MRFs
<https://papers.nips.cc/paper/9687-amortized-bethe-free-energy-minimization-for-learning-mrfs>

AUTHORS: Sam Wiseman, Yoon Kim

HIGHLIGHT: We propose to learn deep undirected graphical models (i.e., MRFs) with a non-ELBO objective for which we can calculate exact gradients.

1393, **TITLE:** Small ReLU networks are powerful memorizers: a tight analysis of memorization capacity
<https://papers.nips.cc/paper/9688-small-relu-networks-are-powerful-memorizers-a-tight-analysis-of-memorization-capacity>
AUTHORS: Chulhee Yun, Suvrit Sra, Ali Jadbabaie
HIGHLIGHT: We study finite sample expressivity, i.e., memorization power of ReLU networks.

1394, **TITLE:** Legendre Memory Units: Continuous-Time Representation in Recurrent Neural Networks
<https://papers.nips.cc/paper/9689-legendre-memory-units-continuous-time-representation-in-recurrent-neural-networks>
AUTHORS: Aaron Voelker, Ivana Kajić, Chris Eliasmith
HIGHLIGHT: We propose a novel memory cell for recurrent neural networks that dynamically maintains information across long windows of time using relatively few resources.

1395, **TITLE:** Exact Combinatorial Optimization with Graph Convolutional Neural Networks
<https://papers.nips.cc/paper/9690-exact-combinatorial-optimization-with-graph-convolutional-neural-networks>
AUTHORS: Maxime Gasse, Didier Chetelat, Nicola Ferroni, Laurent Charlin, Andrea Lodi
HIGHLIGHT: We propose a new graph convolutional neural network model for learning branch-and-bound variable selection policies, which leverages the natural variable-constraint bipartite graph representation of mixed-integer linear programs.

1396, **TITLE:** Fast structure learning with modular regularization
<https://papers.nips.cc/paper/9691-fast-structure-learning-with-modular-regularization>
AUTHORS: Greg Ver Steeg, Hrayr Harutyunyan, Daniel Moyer, Aram Galstyan
HIGHLIGHT: We introduce a novel method that leverages a newly discovered connection between information-theoretic measures and structured latent factor models to derive an optimization objective which encourages modular structures where each observed variable has a single latent parent.

1397, **TITLE:** Wasserstein Dependency Measure for Representation Learning
<https://papers.nips.cc/paper/9692-wasserstein-dependency-measure-for-representation-learning>
AUTHORS: Sherjil Ozair, Corey Lynch, Yoshua Bengio, Aaron van den Oord, Sergey Levine, Pierre Sermanet
HIGHLIGHT: In this work, we empirically demonstrate that mutual information-based representation learning approaches do fail to learn complete representations on a number of designed and real-world tasks.

1398, **TITLE:** TAB-VCR: Tags and Attributes based VCR Baselines
<https://papers.nips.cc/paper/9693-tab-vcr-tags-and-attributes-based-vcr-baselines>
AUTHORS: Jingxiang Lin, Unnat Jain, Alexander Schwing
HIGHLIGHT: Here we show that a much simpler model obtained by ablating and pruning the existing intricate baseline can perform better with half the number of trainable parameters.

1399, **TITLE:** Universality and individuality in neural dynamics across large populations of recurrent networks
<https://papers.nips.cc/paper/9694-universality-and-individuality-in-neural-dynamics-across-large-populations-of-recurrent-networks>
AUTHORS: Niru Maheswaranathan, Alex Williams, Matthew Golub, Surya Ganguli, David Sussillo
HIGHLIGHT: To address these foundational questions, we study populations of thousands of networks of commonly used RNN architectures trained to solve neuroscientifically motivated tasks and characterize their low-dimensional dynamics via CCA and nonlinear dynamical systems analysis.

1400, **TITLE:** End-to-End Learning on 3D Protein Structure for Interface Prediction
<https://papers.nips.cc/paper/9695-end-to-end-learning-on-3d-protein-structure-for-interface-prediction>
AUTHORS: Raphael Townshend, Rishi Bedi, Patricia Suriana, Ron Dror
HIGHLIGHT: To address this question, we focused on a central problem in biology: predicting how proteins interact with one another—that is, which surfaces of one protein bind to those of another protein.

1401, **TITLE:** A Family of Robust Stochastic Operators for Reinforcement Learning
<https://papers.nips.cc/paper/9696-a-family-of-robust-stochastic-operators-for-reinforcement-learning>
AUTHORS: Yingdong Lu, Mark Squillante, Chai Wah Wu
HIGHLIGHT: We consider a new family of stochastic operators for reinforcement learning with the goal of alleviating negative effects and becoming more robust to approximation or estimation errors.

1402, **TITLE:** Using Self-Supervised Learning Can Improve Model Robustness and Uncertainty
<https://papers.nips.cc/paper/9697-using-self-supervised-learning-can-improve-model-robustness-and-uncertainty>

AUTHORS: Dan Hendrycks, Mantas Mazeika, Saurav Kadavath, Dawn Song
HIGHLIGHT: We find that self-supervision can benefit robustness in a variety of ways, including robustness to adversarial examples, label corruption, and common input corruptions.

1403, TITLE: Inherent Tradeoffs in Learning Fair Representations
<https://papers.nips.cc/paper/9698-inherent-tradeoffs-in-learning-fair-representations>
AUTHORS: Han Zhao, Geoff Gordon
HIGHLIGHT: In this paper, through the lens of information theory, we provide the first result that quantitatively characterizes the tradeoff between demographic parity and the joint utility across different population groups.

1404, TITLE: Are deep ResNets provably better than linear predictors?
<https://papers.nips.cc/paper/9699-are-deep-resnets-provably-better-than-linear-predictors>
AUTHORS: Chulhee Yun, Suvrit Sra, Ali Jadbabaie
HIGHLIGHT: Our main theorem on deep ResNets shows under simple geometric conditions that, any critical point in the optimization landscape is either (i) at least as good as the best linear predictor; or (ii) the Hessian at this critical point has a strictly negative eigenvalue.

1405, TITLE: Reverse engineering recurrent networks for sentiment classification reveals line attractor dynamics
<https://papers.nips.cc/paper/9700-reverse-engineering-recurrent-networks-for-sentiment-classification-reveals-line-attractor-dynamics>
AUTHORS: Niru Maheswaranathan, Alex Williams, Matthew Golub, Surya Ganguli, David Sussillo
HIGHLIGHT: In this work, we use tools from dynamical systems analysis to reverse engineer recurrent networks trained to perform sentiment classification, a foundational natural language processing task.

1406, TITLE: BehaveNet: nonlinear embedding and Bayesian neural decoding of behavioral videos
<https://papers.nips.cc/paper/9701-behavenet-nonlinear-embedding-and-bayesian-neural-decoding-of-behavioral-videos>
AUTHORS: Eleanor Batty, Matthew Whiteaway, Shreya Saxena, Dan Biderman, Taiga Abe, Simon Musall, Winthrop Gillis, Jeffrey Markowitz, Anne Churchland, John P. Cunningham, Sandeep R. Datta, Scott Linderman, Liam Paninski
HIGHLIGHT: Here we introduce a probabilistic framework for the analysis of behavioral video and neural activity.

1407, TITLE: Variational Mixture-of-Experts Autoencoders for Multi-Modal Deep Generative Models
<https://papers.nips.cc/paper/9702-variational-mixture-of-experts-autoencoders-for-multi-modal-deep-generative-models>
AUTHORS: Yuge Shi, Siddharth N, Brooks Paige, Philip Torr
HIGHLIGHT: In this work, we characterise successful learning of such models as the fulfilment of four criteria: i) implicit latent decomposition into shared and private subspaces, ii) coherent joint generation over all modalities, iii) coherent cross-generation across individual modalities, and iv) improved model learning for individual modalities through multi-modal integration.

1408, TITLE: Gradient-based Adaptive Markov Chain Monte Carlo
<https://papers.nips.cc/paper/9703-gradient-based-adaptive-markov-chain-monte-carlo>
AUTHORS: Michalis Titsias, Petros Dellaportas
HIGHLIGHT: We introduce a gradient-based learning method to automatically adapt Markov chain Monte Carlo (MCMC) proposal distributions to intractable targets.

1409, TITLE: On the Transfer of Inductive Bias from Simulation to the Real World: a New Disentanglement Dataset
<https://papers.nips.cc/paper/9704-on-the-transfer-of-inductive-bias-from-simulation-to-the-real-world-a-new-disentanglement-dataset>
AUTHORS: Muhammad Waleed Gondal, Manuel Wuthrich, Djordje Miladinovic, Francesco Locatello, Martin Breidt, Valentin Volchkov, Joel Akpo, Olivier Bachem, Bernhard Schölkopf, Stefan Bauer
HIGHLIGHT: In this paper, we propose a novel data-set which consists of over 1 million images of physical 3D objects with seven factors of variation, such as object color, shape, size and position.

1410, TITLE: Imitation-Projected Programmatic Reinforcement Learning
<https://papers.nips.cc/paper/9705-imitation-projected-programmatic-reinforcement-learning>
AUTHORS: Abhinav Verma, Hoang Le, Yisong Yue, Swarat Chaudhuri
HIGHLIGHT: We study the problem of programmatic reinforcement learning, in which policies are represented as short programs in a symbolic language.

1411, TITLE: Learning Data Manipulation for Augmentation and Weighting
<https://papers.nips.cc/paper/9706-learning-data-manipulation-for-augmentation-and-weighting>
AUTHORS: Zhiting Hu, Bowen Tan, Russ R. Salakhutdinov, Tom M. Mitchell, Eric P. Xing
HIGHLIGHT: In this work, we propose a new method that supports learning different manipulation schemes with the same gradient-based algorithm.

1412, TITLE: Exploring Algorithmic Fairness in Robust Graph Covering Problems
<https://papers.nips.cc/paper/9707-exploring-algorithmic-fairness-in-robust-graph-covering-problems>
AUTHORS: Aida Rahmattalabi, Phebe Vayanos, Anthony Fulginiti, Eric Rice, Bryan Wilder, Amulya Yadav, Milind Tambe
HIGHLIGHT: To remediate this issue, we propose a novel formulation of the robust covering problem with fairness constraints and a tractable approximation scheme applicable to real world instances.

1413, TITLE: Abstraction based Output Range Analysis for Neural Networks
<https://papers.nips.cc/paper/9708-abstraction-based-output-range-analysis-for-neural-networks>
AUTHORS: Pavithra Prabhakar, Zahra Rahimi Afzal
HIGHLIGHT: In this paper, we consider the problem of output range analysis for feed-forward neural networks.

1414, TITLE: Space and Time Efficient Kernel Density Estimation in High Dimensions
<https://papers.nips.cc/paper/9709-space-and-time-efficient-kernel-density-estimation-in-high-dimensions>
AUTHORS: Arturs Backurs, Piotr Indyk, Tal Wagner
HIGHLIGHT: In this work, we present an improvement to their framework that retains the same query time, while requiring only linear space and linear preprocessing time.

1415, TITLE: PIDForest: Anomaly Detection via Partial Identification
<https://papers.nips.cc/paper/9710-pidforest-anomaly-detection-via-partial-identification>
AUTHORS: Parikshit Gopalan, Vatsal Sharan, Udi Wieder
HIGHLIGHT: We propose a framework called Partial Identification which captures the intuition that anomalies are easy to distinguish from the overwhelming majority of points by relatively few attribute values.

1416, TITLE: Generative Models for Graph-Based Protein Design
<https://papers.nips.cc/paper/9711-generative-models-for-graph-based-protein-design>
AUTHORS: John Ingraham, Vikas Garg, Regina Barzilay, Tommi Jaakkola
HIGHLIGHT: We develop relational language models for protein sequences that directly condition on a graph specification of the target structure.

1417, TITLE: The Geometry of Deep Networks: Power Diagram Subdivision
<https://papers.nips.cc/paper/9712-the-geometry-of-deep-networks-power-diagram-subdivision>
AUTHORS: Randall Balestriero, Romain Cosentino, Behnaam Aazhang, Richard Baraniuk
HIGHLIGHT: We study the geometry of deep (neural) networks (DNs) with piecewise affine and convex nonlinearities.

1418, TITLE: Approximate Feature Collisions in Neural Nets
<https://papers.nips.cc/paper/9713-approximate-feature-collisions-in-neural-nets>
AUTHORS: Ke Li, Tianhao Zhang, Jitendra Malik
HIGHLIGHT: In this paper, we show the opposite: neural nets could be surprisingly insensitive to adversarially chosen changes of large magnitude.

1419, TITLE: Ease-of-Teaching and Language Structure from Emergent Communication
<https://papers.nips.cc/paper/9714-ease-of-teaching-and-language-structure-from-emergent-communication>
AUTHORS: Fushan Li, Michael Bowling
HIGHLIGHT: By introducing new agents periodically to replace old ones, sequentially and within a population, we explore such a new pressure - ease of teaching - and show its impact on the structure of the resulting language.

1420, TITLE: Generalization in multitask deep neural classifiers: a statistical physics approach
<https://papers.nips.cc/paper/9715-generalization-in-multitask-deep-neural-classifiers-a-statistical-physics-approach>
AUTHORS: Anthony Ndirango, Tyler Lee
HIGHLIGHT: We develop an analytic theory of the nonlinear dynamics of generalization of deep neural networks trained to solve classification tasks using softmax outputs and cross-entropy loss, addressing both single task and multitask settings.

1421, TITLE: Optimistic Distributionally Robust Optimization for Nonparametric Likelihood Approximation
<https://papers.nips.cc/paper/9716-optimistic-distributionally-robust-optimization-for-nonparametric-likelihood-approximation>
AUTHORS: Viet Anh Nguyen, Soroosh Shafieezadeh Abadeh, Man-Chung Yue, Daniel Kuhn, Wolfram Wiesemann
HIGHLIGHT: In this paper, we propose a non-parametric approximation of the likelihood that identifies a probability measure which lies in the neighborhood of the nominal measure and that maximizes the probability of observing the given sample point.

1422, TITLE: On Relating Explanations and Adversarial Examples

<https://papers.nips.cc/paper/9717-on-relating-explanations-and-adversarial-examples>

AUTHORS: Alexey Ignatiev, Nina Narodytska, Joao Marques-Silva

HIGHLIGHT: This paper demonstrates that explanations and adversarial examples are related by a generalized form of hitting set duality, which extends earlier work on hitting set duality observed in model-based diagnosis and knowledge compilation.

1423, TITLE: On the equivalence between graph isomorphism testing and function approximation with GNNs

<https://papers.nips.cc/paper/9718-on-the-equivalence-between-graph-isomorphism-testing-and-function-approximation-with-gnns>

AUTHORS: Zhengdao Chen, Soledad Villar, Lei Chen, Joan Bruna

HIGHLIGHT: Our work connects these two perspectives and proves their equivalence.

1424, TITLE: Surround Modulation: A Bio-inspired Connectivity Structure for Convolutional Neural Networks

<https://papers.nips.cc/paper/9719-surround-modulation-a-bio-inspired-connectivity-structure-for-convolutional-neural-networks>

AUTHORS: Hosein Hasani, Mahdieh Soleymani, Hamid Aghajan

HIGHLIGHT: Inspired by the notion of surround modulation, we designed new excitatory-inhibitory connections between a unit and its surrounding units in the convolutional neural network (CNN) to achieve a more biologically plausible network.

1425, TITLE: Self-attention with Functional Time Representation Learning

<https://papers.nips.cc/paper/9720-self-attention-with-functional-time-representation-learning>

AUTHORS: Da Xu, Chuanwei Ruan, Evren Korpeoglu, Sushant Kumar, Kannan Achan

HIGHLIGHT: We propose several models to learn the functional time representation and the interactions with event representation.

1426, TITLE: Re-randomized Densification for One Permutation Hashing and Bin-wise Consistent Weighted Sampling

<https://papers.nips.cc/paper/9721-re-randomized-densification-for-one-permutation-hashing-and-bin-wise-consistent-weighted-sampling>

AUTHORS: Ping Li, Xiaoyun Li, Cun-Hui Zhang

HIGHLIGHT: In this paper, we propose a strategy named "re-randomization" in the process of densification that could achieve the smallest variance among all densification schemes.

1427, TITLE: Enabling hyperparameter optimization in sequential autoencoders for spiking neural data

<https://papers.nips.cc/paper/9722-enabling-hyperparameter-optimization-in-sequential-autoencoders-for-spiking-neural-data>

AUTHORS: Mohammad Reza Keshtkaran, Chethan Pandarinath

HIGHLIGHT: We develop and test two potential solutions: an alternate validation method ("sample validation") and a novel regularization method ("coordinated dropout"). These innovations prevent overfitting quite effectively, and allow us to test whether SAEs can achieve good performance on limited data through large-scale HP optimization.