

- 1, TITLE: Scan Order in Gibbs Sampling: Models in Which it Matters and Bounds on How Much  
<https://papers.nips.cc/paper/6589-scan-order-in-gibbs-sampling-models-in-which-it-matters-and-bounds-on-how-much>  
AUTHORS: Bryan D. He, Christopher M. De Sa, Ioannis Mitliagkas, Christopher R?  
HIGHLIGHT: To prove these relative bounds, we introduce a method of augmenting the state space to study systematic scan using conductance.
  
- 2, TITLE: Deep ADMM-Net for Compressive Sensing MRI  
<https://papers.nips.cc/paper/6406-deep-admm-net-for-compressive-sensing-mri>  
AUTHORS: yan yang, Jian Sun, Huibin Li, Zongben Xu  
HIGHLIGHT: To improve the current MRI system in reconstruction accuracy and computational speed, in this paper, we propose a novel deep architecture, dubbed ADMM-Net.
  
- 3, TITLE: A scaled Bregman theorem with applications  
<https://papers.nips.cc/paper/6176-a-scaled-bregman-theorem-with-applications>  
AUTHORS: Richard Nock, Aditya Menon, Cheng Soon Ong  
HIGHLIGHT: We present a new theorem which shows that "Bregman distortions" (employing a potentially non-convex generator) may be exactly re-written as a scaled Bregman divergence computed over transformed data.
  
- 4, TITLE: Swapout: Learning an ensemble of deep architectures  
<https://papers.nips.cc/paper/6205-swapout-learning-an-ensemble-of-deep-architectures>  
AUTHORS: Saurabh Singh, Derek Hoiem, David Forsyth  
HIGHLIGHT: We describe Swapout, a new stochastic training method, that outperforms ResNets of identical network structure yielding impressive results on CIFAR-10 and CIFAR-100.
  
- 5, TITLE: On Regularizing Rademacher Observation Losses  
<https://papers.nips.cc/paper/6572-on-regularizing-rademacher-observation-losses>  
AUTHORS: Richard Nock  
HIGHLIGHT: This is our first contribution: we show how they can be fit into the same theory for the equivalence between example and rado losses.
  
- 6, TITLE: Without-Replacement Sampling for Stochastic Gradient Methods  
<https://papers.nips.cc/paper/6245-without-replacement-sampling-for-stochastic-gradient-methods>  
AUTHORS: Ohad Shamir  
HIGHLIGHT: In this paper, we provide competitive convergence guarantees for without-replacement sampling under several scenarios, focusing on the natural regime of few passes over the data.
  
- 7, TITLE: Fast and Provably Good Seedings for k-Means  
<https://papers.nips.cc/paper/6478-fast-and-provably-good-seedings-for-k-means>  
AUTHORS: Olivier Bachem, Mario Lucic, Hamed Hassani, Andreas Krause  
HIGHLIGHT: We propose a simple yet fast seeding algorithm that produces \*provably\* good clusterings even \*without assumptions\* on the data.
  
- 8, TITLE: Unsupervised Learning for Physical Interaction through Video Prediction  
<https://papers.nips.cc/paper/6161-unsupervised-learning-for-physical-interaction-through-video-prediction>  
AUTHORS: Chelsea Finn, Ian Goodfellow, Sergey Levine  
HIGHLIGHT: To learn about physical object motion without labels, we develop an action-conditioned video prediction model that explicitly models pixel motion, by predicting a distribution over pixel motion from previous frames. To explore video prediction for real-world interactive agents, we also introduce a dataset of 59,000 robot interactions involving pushing motions, including a test set with novel objects.
  
- 9, TITLE: High-Rank Matrix Completion and Clustering under Self-Expressive Models  
<https://papers.nips.cc/paper/6357-high-rank-matrix-completion-and-clustering-under-self-expressive-models>  
AUTHORS: Ehsan Elhamifar  
HIGHLIGHT: We propose efficient algorithms for simultaneous clustering and completion of incomplete high-dimensional data that lie in a union of low-dimensional subspaces.
  
- 10, TITLE: Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling  
<https://papers.nips.cc/paper/6096-learning-a-probabilistic-latent-space-of-object-shapes-via-3d-generative-adversarial-modeling>  
AUTHORS: Jiajun Wu, Chengkai Zhang, Tianfan Xue, Bill Freeman, Josh Tenenbaum  
HIGHLIGHT: We propose a novel framework, namely 3D Generative Adversarial Network (3D-GAN), which generates 3D objects from a probabilistic space by leveraging recent advances in volumetric convolutional networks and generative adversarial nets.

- 11, TITLE: Visual Dynamics: Probabilistic Future Frame Synthesis via Cross Convolutional Networks  
<https://papers.nips.cc/paper/6552-visual-dynamics-probabilistic-future-frame-synthesis-via-cross-convolutional-networks>  
AUTHORS: Tianfan Xue, Jiajun Wu, Katherine Bouman, Bill Freeman  
HIGHLIGHT: In contrast to traditional methods, which have tackled this problem in a deterministic or non-parametric way, we propose a novel approach which models future frames in a probabilistic manner.
- 12, TITLE: Human Decision-Making under Limited Time  
<https://papers.nips.cc/paper/6249-human-decision-making-under-limited-time>  
AUTHORS: Pedro A. Ortega, Alan A. Stocker  
HIGHLIGHT: Abstract Subjective expected utility theory assumes that decision-makers possess unlimited computational resources to reason about their choices; however, virtually all decisions in everyday life are made under resource constraints---i.e. decision-makers are bounded in their rationality.
- 13, TITLE: Incremental Boosting Convolutional Neural Network for Facial Action Unit Recognition  
<https://papers.nips.cc/paper/6258-incremental-boosting-convolutional-neural-network-for-facial-action-unit-recognition>  
AUTHORS: Shizhong Han, Zibo Meng, AHMED-SHEHAB KHAN, Yan Tong  
HIGHLIGHT: We proposed a novel Incremental Boosting CNN (IB-CNN) to integrate boosting into the CNN via an incremental boosting layer that selects discriminative neurons from the lower layer and is incrementally updated on successive mini-batches.
- 14, TITLE: Natural-Parameter Networks: A Class of Probabilistic Neural Networks  
<https://papers.nips.cc/paper/6279-natural-parameter-networks-a-class-of-probabilistic-neural-networks>  
AUTHORS: Hao Wang, Xingjian SHI, Dit-Yan Yeung  
HIGHLIGHT: To address these problems, we propose a class of probabilistic neural networks, dubbed natural-parameter networks (NPN), as a novel and lightweight Bayesian treatment of NN.
- 15, TITLE: Tree-Structured Reinforcement Learning for Sequential Object Localization  
<https://papers.nips.cc/paper/6532-tree-structured-reinforcement-learning-for-sequential-object-localization>  
AUTHORS: Zequn Jie, Xiaodan Liang, Jiashi Feng, Xiaojie Jin, Wen Lu, Shuicheng Yan  
HIGHLIGHT: To incorporate global interdependency between objects into object localization, we propose an effective Tree-structured Reinforcement Learning (Tree-RL) approach to sequentially search for objects by fully exploiting both the current observation and historical search paths.
- 16, TITLE: Unsupervised Domain Adaptation with Residual Transfer Networks  
<https://papers.nips.cc/paper/6110-unsupervised-domain-adaptation-with-residual-transfer-networks>  
AUTHORS: Mingsheng Long, Han Zhu, Jianmin Wang, Michael I. Jordan  
HIGHLIGHT: In this paper, we propose a new approach to domain adaptation in deep networks that can jointly learn adaptive classifiers and transferable features from labeled data in the source domain and unlabeled data in the target domain.
- 17, TITLE: Verification Based Solution for Structured MAB Problems  
<https://papers.nips.cc/paper/6525-verification-based-solution-for-structured-mab-problems>  
AUTHORS: Zohar S. Karnin  
HIGHLIGHT: We consider the problem of finding the best arm in a stochastic Mutli-armed Bandit (MAB) game and propose a general framework based on verification that applies to multiple well-motivated generalizations of the classic MAB problem.
- 18, TITLE: Minimizing Regret on Reflexive Banach Spaces and Nash Equilibria in Continuous Zero-Sum Games  
<https://papers.nips.cc/paper/6216-minimizing-regret-on-reflexive-banach-spaces-and-nash-equilibria-in-continuous-zero-sum-games>  
AUTHORS: Maximilian Balandat, Walid Krichene, Claire Tomlin, Alexandre Bayen  
HIGHLIGHT: Our goal is to minimize regret, defined as the gap between the realized reward and the reward of the best fixed action in hindsight.
- 19, TITLE: Linear dynamical neural population models through nonlinear embeddings  
<https://papers.nips.cc/paper/6430-linear-dynamical-neural-population-models-through-nonlinear-embeddings>  
AUTHORS: Yuanjun Gao, Evan W. Archer, Liam Paninski, John P. Cunningham  
HIGHLIGHT: Here, we propose fLDS, a general class of nonlinear generative models that permits the firing rate of each neuron to vary as an arbitrary smooth function of a latent, linear dynamical state.
- 20, TITLE: SURGE: Surface Regularized Geometry Estimation from a Single Image  
<https://papers.nips.cc/paper/6502-surge-surface-regularized-geometry-estimation-from-a-single-image>

AUTHORS: Peng Wang, Xiaohui Shen, Bryan Russell, Scott Cohen, Brian Price, Alan L. Yuille  
HIGHLIGHT: This paper introduces an approach to regularize 2.5D surface normal and depth predictions at each pixel given a single input image.

21, TITLE: Interpretable Distribution Features with Maximum Testing Power  
<https://papers.nips.cc/paper/6148-interpretability-distribution-features-with-maximum-testing-power>  
AUTHORS: Wittawat Jitkrittum, Zoltan Szabó, Kacper P. Chwialkowski, Arthur Gretton  
HIGHLIGHT: Interpretable Distribution Features with Maximum Testing Power

22, TITLE: Sorting out typicality with the inverse moment matrix SOS polynomial  
<https://papers.nips.cc/paper/6495-sorting-out-typicality-with-the-inverse-moment-matrix-sos-polynomial>  
AUTHORS: Edouard Pauwels, Jean B. Lasserre  
HIGHLIGHT: We study a surprising phenomenon related to the representation of a cloud of data points using polynomials.

23, TITLE: Multi-armed Bandits: Competing with Optimal Sequences  
<https://papers.nips.cc/paper/6341-multi-armed-bandits-competing-with-optimal-sequences>  
AUTHORS: Zohar S. Karmin, Oren Anava  
HIGHLIGHT: We consider sequential decision making problem in the adversarial setting, where regret is measured with respect to the optimal sequence of actions and the feedback adheres the bandit setting.

24, TITLE: Multivariate tests of association based on univariate tests  
<https://papers.nips.cc/paper/6220-multivariate-tests-of-association-based-on-univariate-tests>  
AUTHORS: Ruth Heller, Yair Heller  
HIGHLIGHT: For testing two vector random variables for independence, we propose testing whether the distance of one vector from an arbitrary center point is independent from the distance of the other vector from another arbitrary center point by a univariate test.

25, TITLE: Learning What and Where to Draw  
<https://papers.nips.cc/paper/6111-learning-what-and-where-to-draw>  
AUTHORS: Scott E. Reed, Zeynep Akata, Santosh Mohan, Samuel Tenka, Bernt Schiele, Honglak Lee  
HIGHLIGHT: We propose a new model, the Generative Adversarial What-Where Network (GAWWN), that synthesizes images given instructions describing what content to draw in which location.

26, TITLE: The Sound of APALM Clapping: Faster Nonsmooth Nonconvex Optimization with Stochastic Asynchronous PALM  
<https://papers.nips.cc/paper/6428-the-sound-of-apalm-clapping-faster-nonsmooth-nonconvex-optimization-with-stochastic-asynchronous-palm>  
AUTHORS: Damek Davis, Brent Edmunds, Madeleine Udell  
HIGHLIGHT: We introduce the Stochastic Asynchronous Proximal Alternating Linearized Minimization (SAPALM) method, a block coordinate stochastic proximal-gradient method for solving nonconvex, nonsmooth optimization problems.

27, TITLE: Integrated perception with recurrent multi-task neural networks  
<https://papers.nips.cc/paper/6393-integrated-perception-with-recurrent-multi-task-neural-networks>  
AUTHORS: Hakan Bilen, Andrea Vedaldi  
HIGHLIGHT: In order to capture some of these advantages in machine perception, we ask two questions: whether deep neural networks can learn universal image representations, useful not only for a single task but for all of them, and how the solutions to the different tasks can be integrated in this framework.

28, TITLE: Learning from Small Sample Sets by Combining Unsupervised Meta-Training with CNNs  
<https://papers.nips.cc/paper/6408-learning-from-small-sample-sets-by-combining-unsupervised-meta-training-with-cnns>  
AUTHORS: Yu-Xiong Wang, Martial Hebert  
HIGHLIGHT: Inspired by the transferability properties of CNNs, we introduce an additional unsupervised meta-training stage that exposes multiple top layer units to a large amount of unlabeled real-world images.

29, TITLE: CNNpack: Packing Convolutional Neural Networks in the Frequency Domain  
<https://papers.nips.cc/paper/6390-cnnpack-packing-convolutional-neural-networks-in-the-frequency-domain>  
AUTHORS: Yunhe Wang, Chang Xu, Shan You, Dacheng Tao, Chao Xu  
HIGHLIGHT: Here we present an effective CNN compression approach in the frequency domain, which focuses not only on smaller weights but on all the weights and their underlying connections.

- 30, TITLE: Cooperative Graphical Models  
<https://papers.nips.cc/paper/6122-cooperative-graphical-models>  
AUTHORS: Josip Djolonga, Stefanie Jegelka, Sebastian Tschiatschek, Andreas Krause  
HIGHLIGHT: Yet, this family retains structure, which we carefully exploit for efficient inference techniques.
- 31, TITLE: f-GAN: Training Generative Neural Samplers using Variational Divergence Minimization  
<https://papers.nips.cc/paper/6066-f-gan-training-generative-neural-samplers-using-variational-divergence-minimization>  
AUTHORS: Sebastian Nowozin, Botond Cseke, Ryota Tomioka  
HIGHLIGHT: The generative-adversarial training method allows to train such models through the use of an auxiliary discriminative neural network.
- 32, TITLE: Bayesian Optimization for Probabilistic Programs  
<https://papers.nips.cc/paper/6421-bayesian-optimization-for-probabilistic-programs>  
AUTHORS: Tom Rainforth, Tuan Anh Le, Jan-Willem van de Meent, Michael A. Osborne, Frank Wood  
HIGHLIGHT: We present the first general purpose framework for marginal maximum a posteriori estimation of probabilistic program variables.
- 33, TITLE: Hierarchical Question-Image Co-Attention for Visual Question Answering  
<https://papers.nips.cc/paper/6202-hierarchical-question-image-co-attention-for-visual-question-answering>  
AUTHORS: Jiasen Lu, Jianwei Yang, Dhruv Batra, Devi Parikh  
HIGHLIGHT: We present a novel co-attention model for VQA that jointly reasons about image and question attention.
- 34, TITLE: Optimal Sparse Linear Encoders and Sparse PCA  
<https://papers.nips.cc/paper/6252-optimal-sparse-linear-encoders-and-sparse-pca>  
AUTHORS: Malik Magdon-Ismail, Christos Boutsidis  
HIGHLIGHT: Optimal Sparse Linear Encoders and Sparse PCA
- 35, TITLE: FPNN: Field Probing Neural Networks for 3D Data  
<https://papers.nips.cc/paper/6416-fpnn-field-probing-neural-networks-for-3d-data>  
AUTHORS: Yangyan Li, Soeren Pirk, Hao Su, Charles R. Qi, Leonidas J. Guibas  
HIGHLIGHT: In this work, we represent 3D spaces as volumetric fields, and propose a novel design that employs field probing filters to efficiently extract features from them.
- 36, TITLE: CRF-CNN: Modeling Structured Information in Human Pose Estimation  
<https://papers.nips.cc/paper/6278-crf-cnn-modeling-structured-information-in-human-pose-estimation>  
AUTHORS: Xiao Chu, Wanli Ouyang, hongsheng Li, Xiaogang Wang  
HIGHLIGHT: In this paper, we propose a CRF-CNN framework which can simultaneously model structural information in both output and hidden feature layers in a probabilistic way, and it is applied to human pose estimation.
- 37, TITLE: Fairness in Learning: Classic and Contextual Bandits  
<https://papers.nips.cc/paper/6355-fairness-in-learning-classic-and-contextual-bandits>  
AUTHORS: Matthew Joseph, Michael Kearns, Jamie H. Morgenstern, Aaron Roth  
HIGHLIGHT: We introduce the study of fairness in multi-armed bandit problems.
- 38, TITLE: Joint M-Best-Diverse Labelings as a Parametric Submodular Minimization  
<https://papers.nips.cc/paper/6396-joint-m-best-diverse-labelings-as-a-parametric-submodular-minimization>  
AUTHORS: Alexander Kirillov, Alexander Shekhovtsov, Carsten Rother, Bogdan Savchynskyy  
HIGHLIGHT: As the main contribution of this work we establish a close relationship between diversity with submodular energies and the parametric submodular minimization.
- 39, TITLE: Domain Separation Networks  
<https://papers.nips.cc/paper/6254-domain-separation-networks>  
AUTHORS: Konstantinos Bousmalis, George Trigeorgis, Nathan Silberman, Dilip Krishnan, Dumitru Erhan  
HIGHLIGHT: Existing approaches focus either on mapping representations from one domain to the other, or on learning to extract features that are invariant to the domain from which they were extracted.
- 40, TITLE: DISCO Nets : DISsimilarity COefficients Networks  
<https://papers.nips.cc/paper/6143-disco-nets-dissimilarity-coefficients-networks>  
AUTHORS: Diane Bouchacourt, Pawan K. Mudigonda, Sebastian Nowozin

HIGHLIGHT: We present a new type of probabilistic model which we call DISsimilarity COefficient Networks (DISCO Nets).

41, TITLE: Multimodal Residual Learning for Visual QA  
<https://papers.nips.cc/paper/6446-multimodal-residual-learning-for-visual-qa>  
AUTHORS: Jin-Hwa Kim, Sang-Woo Lee, Donghyun Kwak, Min-Oh Heo, Jeonghee Kim, Jung-Woo Ha, Byoung-Tak Zhang  
HIGHLIGHT: We present Multimodal Residual Networks (MRN) for the multimodal residual learning of visual question-answering, which extends the idea of the deep residual learning.

42, TITLE: CMA-ES with Optimal Covariance Update and Storage Complexity  
<https://papers.nips.cc/paper/6457-cma-es-with-optimal-covariance-update-and-storage-complexity>  
AUTHORS: Oswin Krause, Dac Rodriguez Arbon, Christian Igel  
HIGHLIGHT: We propose a numerically stable quadratic-time covariance matrix update scheme with minimal memory requirements based on maintaining triangular Cholesky factors.

43, TITLE: R-FCN: Object Detection via Region-based Fully Convolutional Networks  
<https://papers.nips.cc/paper/6465-r-fcn-object-detection-via-region-based-fully-convolutional-networks>  
AUTHORS: Jifeng Dai, Yi Li, Kaiming He, Jian Sun  
HIGHLIGHT: We present region-based, fully convolutional networks for accurate and efficient object detection.

44, TITLE: GAP Safe Screening Rules for Sparse-Group Lasso  
<https://papers.nips.cc/paper/6405-gap-safe-screening-rules-for-sparse-group-lasso>  
AUTHORS: Eugene Ndiaye, Olivier Fercoq, Alexandre Gramfort, Joseph Salmon  
HIGHLIGHT: We propose the first (provably) safe screening rules for Sparse-Group Lasso, i.e., rules that allow to discard early in the solver features/groups that are inactive at optimal solution.

45, TITLE: Learning and Forecasting Opinion Dynamics in Social Networks  
<https://papers.nips.cc/paper/6193-learning-and-forecasting-opinion-dynamics-in-social-networks>  
AUTHORS: Abir De, Isabel Valera, Niloy Ganguly, Sourangshu Bhattacharya, Manuel Gomez Rodriguez  
HIGHLIGHT: In this paper, we introduce SLANT, a probabilistic modeling framework of opinion dynamics, which represents users' opinions over time by means of marked jump diffusion stochastic differential equations, and allows for efficient model simulation and parameter estimation from historical fine grained event data.

46, TITLE: Gradient-based Sampling: An Adaptive Importance Sampling for Least-squares  
<https://papers.nips.cc/paper/6579-gradient-based-sampling-an-adaptive-importance-sampling-for-least-squares>  
AUTHORS: Rong Zhu  
HIGHLIGHT: In this paper we propose an adaptive sampling called the gradient-based sampling which is dependent on both the input data and the output for fast solving of least-square (LS) problems.

47, TITLE: Collaborative Recurrent Autoencoder: Recommend while Learning to Fill in the Blanks  
<https://papers.nips.cc/paper/6163-collaborative-recurrent-autoencoder-recommend-while-learning-to-fill-in-the-blanks>  
AUTHORS: Hao Wang, Xingjian SHI, Dit-Yan Yeung  
HIGHLIGHT: Hybrid methods that utilize both content and rating information are commonly used in many recommender systems.

48, TITLE: Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula  
<https://papers.nips.cc/paper/6380-mutual-information-for-symmetric-rank-one-matrix-estimation-a-proof-of-the-replica-formula>  
AUTHORS: Jean Barbier, Mohamad Dia, Nicolas Macris, Florent Krzakala, Thibault Lesieur, Lenka Zdeborová  
HIGHLIGHT: Here, we show how to rigorously prove the conjectured formula for the symmetric rank-one case.

49, TITLE: A Unified Approach for Learning the Parameters of Sum-Product Networks  
<https://papers.nips.cc/paper/6423-a-unified-approach-for-learning-the-parameters-of-sum-product-networks>  
AUTHORS: Han Zhao, Pascal Poupart, Geoffrey J. Gordon  
HIGHLIGHT: We present a unified approach for learning the parameters of Sum-Product networks (SPNs).

50, TITLE: Training and Evaluating Multimodal Word Embeddings with Large-scale Web Annotated Images  
<https://papers.nips.cc/paper/6590-training-and-evaluating-multimodal-word-embeddings-with-large-scale-web-annotated-images>  
AUTHORS: Junhua Mao, Jiajing Xu, Kevin Jing, Alan L. Yuille

**HIGHLIGHT:** In this paper, we focus on training and evaluating effective word embeddings with both text and visual information.

More specifically, we introduce a large-scale dataset with 300 million sentences describing over 40 million images crawled and downloaded from publicly available Pins (i.e. an image with sentence descriptions uploaded by users) on Pinterest. In addition, we construct an evaluation dataset to directly assess the effectiveness of word embeddings in terms of finding semantically similar or related words and phrases.

51, **TITLE:** Stochastic Online AUC Maximization  
<https://papers.nips.cc/paper/6065-stochastic-online-auc-maximization>  
**AUTHORS:** Yiming Ying, Longyin Wen, Siwei Lyu  
**HIGHLIGHT:** In this work, we propose a new stochastic online algorithm for AUC maximization.

52, **TITLE:** The Generalized Reparameterization Gradient  
<https://papers.nips.cc/paper/6328-the-generalized-reparameterization-gradient>  
**AUTHORS:** Francisco R. Ruiz, Michalis Titsias RC AUEB, David Blei  
**HIGHLIGHT:** In this paper, we introduce the generalized reparameterization gradient, a method that extends the reparameterization gradient to a wider class of variational distributions.

53, **TITLE:** Coupled Generative Adversarial Networks  
<https://papers.nips.cc/paper/6544-coupled-generative-adversarial-networks>  
**AUTHORS:** Ming-Yu Liu, Oncel Tuzel  
**HIGHLIGHT:** We propose the coupled generative adversarial nets (CoGAN) framework for generating pairs of corresponding images in two different domains.

54, **TITLE:** Exponential Family Embeddings  
<https://papers.nips.cc/paper/6571-exponential-family-embeddings>  
**AUTHORS:** Maja Rudolph, Francisco Ruiz, Stephan Mandt, David Blei  
**HIGHLIGHT:** In this paper, we develop exponential family embeddings, which extends the idea of word embeddings to other types of high-dimensional data.

55, **TITLE:** Variational Information Maximization for Feature Selection  
<https://papers.nips.cc/paper/6444-variational-information-maximization-for-feature-selection>  
**AUTHORS:** Shuyang Gao, Greg Ver Steeg, Aram Galstyan  
**HIGHLIGHT:** These bounds define a novel information-theoretic framework for feature selection, which we prove to be optimal under tree graphical models with proper choice of variational distributions.

56, **TITLE:** Operator Variational Inference  
<https://papers.nips.cc/paper/6091-operator-variational-inference>  
**AUTHORS:** Rajesh Ranganath, Dustin Tran, Jaan Altosaar, David Blei  
**HIGHLIGHT:** As one example, we design a variational objective with a Langevin-Stein operator.

57, **TITLE:** Fast learning rates with heavy-tailed losses  
<https://papers.nips.cc/paper/6104-fast-learning-rates-with-heavy-tailed-losses>  
**AUTHORS:** Vu C. Dinh, Lam S. Ho, Binh Nguyen, Duy Nguyen  
**HIGHLIGHT:** To enable such analyses, we introduce two new conditions: (i) the envelope function  $\sup_{f \in \mathcal{F}} \mathbb{E}[\ell(f)]$ , where  $\ell$  is the loss function and  $\mathcal{F}$  is the hypothesis class, exists and is  $L^r$ -integrable, and (ii)  $\ell$  satisfies the multi-scale Bernstein's condition on  $\mathcal{F}$ .

58, **TITLE:** Budgeted stream-based active learning via adaptive submodular maximization  
<https://papers.nips.cc/paper/6038-budgeted-stream-based-active-learning-via-adaptive-submodular-maximization>  
**AUTHORS:** Kaito Fujii, Hisashi Kashima  
**HIGHLIGHT:** In this paper, we propose a new class of utility functions, policy-adaptive submodular functions, and prove this class includes many existing adaptive submodular functions appearing in real world problems.

59, **TITLE:** Learning feed-forward one-shot learners  
<https://papers.nips.cc/paper/6068-learning-feed-forward-one-shot-learners>  
**AUTHORS:** Luca Bertinetto, Jo?o F. Henriques, Jack Valmadre, Philip Torr, Andrea Vedaldi  
**HIGHLIGHT:** In this paper, we propose a method to learn the parameters of a deep model in one shot.

60, **TITLE:** Learning User Perceived Clusters with Feature-Level Supervision

<https://papers.nips.cc/paper/6260-learning-user-perceived-clusters-with-feature-level-supervision>  
AUTHORS: Ting-Yu Cheng, Guiguan Lin, xinyang gong, Kang-Jun Liu, Shan-Hung Wu  
HIGHLIGHT: In this paper, we propose learning from the feature-level supervision.

61, TITLE: Robust Spectral Detection of Global Structures in the Data by Learning a Regularization  
<https://papers.nips.cc/paper/6491-robust-spectral-detection-of-global-structures-in-the-data-by-learning-a-regularization>  
AUTHORS: Pan Zhang  
HIGHLIGHT: In this work, we propose a general method to solve the localization problem by learning a regularization matrix from the localized eigenvectors.

62, TITLE: Residual Networks Behave Like Ensembles of Relatively Shallow Networks  
<https://papers.nips.cc/paper/6556-residual-networks-behave-like-ensembles-of-relatively-shallow-networks>  
AUTHORS: Andreas Veit, Michael J. Wilber, Serge Belongie  
HIGHLIGHT: In this work we propose a novel interpretation of residual networks showing that they can be seen as a collection of many paths of differing length.

63, TITLE: Adversarial Multiclass Classification: A Risk Minimization Perspective  
<https://papers.nips.cc/paper/6088-adversarial-multiclass-classification-a-risk-minimization-perspective>  
AUTHORS: Rizal Fathony, Anqi Liu, Kaiser Asif, Brian Ziebart  
HIGHLIGHT: Recently proposed adversarial classification methods have shown promising results for cost sensitive and multivariate losses.

64, TITLE: Solving Random Systems of Quadratic Equations via Truncated Generalized Gradient Flow  
<https://papers.nips.cc/paper/6061-solving-random-systems-of-quadratic-equations-via-truncated-generalized-gradient-flow>  
AUTHORS: Gang Wang, Georgios Giannakis  
HIGHLIGHT: This paper puts forth a novel algorithm, termed *truncated generalized gradient flow* (TGGF), to solve for  $\mathbf{x} \in \mathbb{R}^n \times \mathbb{C}^n$  a system of  $m$  quadratic equations  $y_i = |\langle \mathbf{a}_i, \mathbf{x} \rangle|^2$ ,  $i=1, 2, \dots, m$ , which even for  $\{\mathbf{a}_i\}_{i=1}^m$  random is known to be *NP-hard* in general.

65, TITLE: Coin Betting and Parameter-Free Online Learning  
<https://papers.nips.cc/paper/6159-coin-betting-and-parameter-free-online-learning>  
AUTHORS: Francesco Orabona, David Pal  
HIGHLIGHT: We present a new intuitive framework to design parameter-free algorithms for both online linear optimization over Hilbert spaces and for learning with expert advice, based on reductions to betting on outcomes of adversarial coins.

66, TITLE: Deep Learning without Poor Local Minima  
<https://papers.nips.cc/paper/6112-deep-learning-without-poor-local-minima>  
AUTHORS: Kenji Kawaguchi  
HIGHLIGHT: In this paper, we prove a conjecture published in 1989 and also partially address an open problem announced at the Conference on Learning Theory (COLT) 2015.

67, TITLE: Testing for Differences in Gaussian Graphical Models: Applications to Brain Connectivity  
<https://papers.nips.cc/paper/6531-testing-for-differences-in-gaussian-graphical-models-applications-to-brain-connectivity>  
AUTHORS: Eugene Belilovsky, Gal Varoquaux, Matthew B. Blaschko  
HIGHLIGHT: Our goal is to identify differences in GGMs known to have similar structure.

68, TITLE: A Constant-Factor Bi-Criteria Approximation Guarantee for  $k$ -means++  
<https://papers.nips.cc/paper/6309-a-constant-factor-bi-criteria-approximation-guarantee-for-k-means>  
AUTHORS: Dennis Wei  
HIGHLIGHT: This paper studies the  $k$ -means++ algorithm for clustering as well as the class of  $\mathcal{D}^{\text{ell}}$  sampling algorithms to which  $k$ -means++ belongs.

69, TITLE: Generating Videos with Scene Dynamics  
<https://papers.nips.cc/paper/6194-generating-videos-with-scene-dynamics>  
AUTHORS: Carl Vondrick, Hamed Pirsiavash, Antonio Torralba  
HIGHLIGHT: We propose a generative adversarial network for video with a spatio-temporal convolutional architecture that untangles the scene's foreground from the background.

- 70, TITLE: Neurally-Guided Procedural Models: Amortized Inference for Procedural Graphics Programs using Neural Networks  
<https://papers.nips.cc/paper/6353-neurally-guided-procedural-models-amortized-inference-for-procedural-graphics-programs-using-neural-networks>  
AUTHORS: Daniel Ritchie, Anna Thomas, Pat Hanrahan, Noah Goodman  
HIGHLIGHT: In this paper, we show how to create procedural models which learn how to satisfy constraints.
- 71, TITLE: A Powerful Generative Model Using Random Weights for the Deep Image Representation  
<https://papers.nips.cc/paper/6568-a-powerful-generative-model-using-random-weights-for-the-deep-image-representation>  
AUTHORS: Kun He, Yan Wang, John Hopcroft  
HIGHLIGHT: To address this issue, we explore new and powerful generative models for three popular deep visualization tasks using untrained, random weight convolutional neural networks.
- 72, TITLE: Optimizing affinity-based binary hashing using auxiliary coordinates  
<https://papers.nips.cc/paper/6281-optimizing-affinity-based-binary-hashing-using-auxiliary-coordinates>  
AUTHORS: Ramin Raziperchikolaei, Miguel A. Carreira-Perpinan  
HIGHLIGHT: We propose a general framework for learning hash functions using affinity-based loss functions that uses auxiliary coordinates.
- 73, TITLE: Double Thompson Sampling for Dueling Bandits  
<https://papers.nips.cc/paper/6157-double-thompson-sampling-for-dueling-bandits>  
AUTHORS: Huasen Wu, Xin Liu  
HIGHLIGHT: In this paper, we propose a Double Thompson Sampling (D-TS) algorithm for dueling bandit problems.
- 74, TITLE: Generating Images with Perceptual Similarity Metrics based on Deep Networks  
<https://papers.nips.cc/paper/6158-generating-images-with-perceptual-similarity-metrics-based-on-deep-networks>  
AUTHORS: Alexey Dosovitskiy, Thomas Brox  
HIGHLIGHT: We propose a class of loss functions, which we call deep perceptual similarity metrics (DeepSiM), allowing to generate sharp high resolution images from compressed abstract representations.
- 75, TITLE: Dynamic Filter Networks  
<https://papers.nips.cc/paper/6578-dynamic-filter-networks>  
AUTHORS: Xu Jia, Bert De Brabandere, Tinne Tuytelaars, Luc V. Gool  
HIGHLIGHT: In contrast, we introduce a new framework, the Dynamic Filter Network, where filters are generated dynamically conditioned on an input.
- 76, TITLE: A Simple Practical Accelerated Method for Finite Sums  
<https://papers.nips.cc/paper/6154-a-simple-practical-accelerated-method-for-finite-sums>  
AUTHORS: Aaron Defazio  
HIGHLIGHT: Abstract We describe a novel optimization method for finite sums (such as empirical risk minimization problems) building on the recently introduced SAGA method.
- 77, TITLE: Barzilai-Borwein Step Size for Stochastic Gradient Descent  
<https://papers.nips.cc/paper/6286-barzilai-borwein-step-size-for-stochastic-gradient-descent>  
AUTHORS: Conghui Tan, Shiqian Ma, Yu-Hong Dai, Yuqiu Qian  
HIGHLIGHT: In this paper, we propose to use the Barzilai-Borwein (BB) method to automatically compute step sizes for SGD and its variant: stochastic variance reduced gradient (SVRG) method, which leads to two algorithms: SGD-BB and SVRG-BB.
- 78, TITLE: On Graph Reconstruction via Empirical Risk Minimization: Fast Learning Rates and Scalability  
<https://papers.nips.cc/paper/6588-on-graph-reconstruction-via-empirical-risk-minimization-fast-learning-rates-and-scalability>  
AUTHORS: Guillaume Papa, Aurélien Bellet, Stephan Clémens  
HIGHLIGHT: Our first contribution is to derive learning rates of order  $O(\log n / n)$  for this problem, significantly improving upon the slow rates of order  $O(1/n)$  established in the seminal work of Biau & Bleakley (2006).
- 79, TITLE: Optimal spectral transportation with application to music transcription  
<https://papers.nips.cc/paper/6479-optimal-spectral-transportation-with-application-to-music-transcription>  
AUTHORS: R?mi Flamary, C?dric F?votte, Nicolas Courty, Valentin Emiya  
HIGHLIGHT: We address these issues by means of optimal transportation and propose a new measure of fit that treats the frequency distributions of energy holistically as opposed to frequency-wise.

- 80, TITLE: Regularized Nonlinear Acceleration  
<https://papers.nips.cc/paper/6267-regularized-nonlinear-acceleration>  
AUTHORS: Damien Scieur, Alexandre d'Aspremont, Francis Bach  
HIGHLIGHT: We describe a convergence acceleration technique for generic optimization problems.
- 81, TITLE: SPALS: Fast Alternating Least Squares via Implicit Leverage Scores Sampling  
<https://papers.nips.cc/paper/6436-spals-fast-alternating-least-squares-via-implicit-leverage-scores-sampling>  
AUTHORS: Dehua Cheng, Richard Peng, Yan Liu, Ioakeim Perros  
HIGHLIGHT: In this paper, we show ways of sampling intermediate steps of alternating minimization algorithms for computing low rank tensor CP decompositions, leading to the sparse alternating least squares (SPALS) method.
- 82, TITLE: Single-Image Depth Perception in the Wild  
<https://papers.nips.cc/paper/6489-single-image-depth-perception-in-the-wild>  
AUTHORS: Weifeng Chen, Zhao Fu, Dawei Yang, Jia Deng  
HIGHLIGHT: We introduce a new dataset "Depth in the Wild" consisting of images in the wild annotated with relative depth between pairs of random points.
- 83, TITLE: Computational and Statistical Tradeoffs in Learning to Rank  
<https://papers.nips.cc/paper/6442-computational-and-statistical-tradeoffs-in-learning-to-rank>  
AUTHORS: Ashish Khetan, Sewoong Oh  
HIGHLIGHT: In the application of learning to rank, we provide a hierarchy of rank-breaking mechanisms ordered by the complexity in thus generated sketch of the data.
- 84, TITLE: Online Convex Optimization with Unconstrained Domains and Losses  
<https://papers.nips.cc/paper/6371-online-convex-optimization-with-unconstrained-domains-and-losses>  
AUTHORS: Ashok Cutkosky, Kwabena A. Boahen  
HIGHLIGHT: We propose an online convex optimization algorithm (RescaledExp) that achieves optimal regret in the unconstrained setting without prior knowledge of any bounds on the loss functions.
- 85, TITLE: An ensemble diversity approach to supervised binary hashing  
<https://papers.nips.cc/paper/6289-an-ensemble-diversity-approach-to-supervised-binary-hashing>  
AUTHORS: Miguel A. Carreira-Perpinan, Ramin Raziperchikolaei  
HIGHLIGHT: We propose a much simpler approach: we train each hash function (or bit) independently from each other, but introduce diversity among them using techniques from classifier ensembles.
- 86, TITLE: Efficient Globally Convergent Stochastic Optimization for Canonical Correlation Analysis  
<https://papers.nips.cc/paper/6459-efficient-globally-convergent-stochastic-optimization-for-canonical-correlation-analysis>  
AUTHORS: Weiran Wang, Jialei Wang, Dan Garber, Dan Garber, Nati Srebro  
HIGHLIGHT: Inspired by the alternating least squares/power iterations formulation of CCA, and the shift-and-invert preconditioning method for PCA, we propose two globally convergent meta-algorithms for CCA, both of which transform the original problem into sequences of least squares problems that need only be solved approximately.
- 87, TITLE: The Power of Adaptivity in Identifying Statistical Alternatives  
<https://papers.nips.cc/paper/6072-the-power-of-adaptivity-in-identifying-statistical-alternatives>  
AUTHORS: Kevin G. Jamieson, Daniel Haas, Benjamin Recht  
HIGHLIGHT: We focus on the most biased coin problem, asking how many total coin flips are required to identify a "heavy" coin from an infinite bag containing both "heavy" coins with mean  $\theta_1$  in  $(0,1)$ , and "light" coins with mean  $\theta_0$  in  $(0,\theta_1)$ , where heavy coins are drawn from the bag with proportion  $\alpha$  in  $(0,1/2)$ .
- 88, TITLE: On Explore-Then-Commit strategies  
<https://papers.nips.cc/paper/6179-on-explore-then-commit-strategies>  
AUTHORS: Aurelien Garivier, Tor Lattimore, Emilie Kaufmann  
HIGHLIGHT: Our objective is to use this simple setting to illustrate that strategies based on an exploration phase (up to a stopping time) followed by exploitation are necessarily suboptimal.
- 89, TITLE: Sublinear Time Orthogonal Tensor Decomposition  
<https://papers.nips.cc/paper/6496-sublinear-time-orthogonal-tensor-decomposition>  
AUTHORS: Zhao Song, David Woodruff, Huan Zhang  
HIGHLIGHT: al., NIPS 2015) gives the fastest known algorithms for orthogonal tensor decomposition with provable guarantees.

- 90, TITLE: DECOrelated feature space partitioning for distributed sparse regression  
<https://papers.nips.cc/paper/6349-decorrelated-feature-space-partitioning-for-distributed-sparse-regression>  
AUTHORS: Xiangyu Wang, David B. Dunson, Chenlei Leng  
HIGHLIGHT: In this paper, we solve these problems through a new embarrassingly parallel framework named DECO for distributed variable selection and parameter estimation.
- 91, TITLE: Deep Alternative Neural Network: Exploring Contexts as Early as Possible for Action Recognition  
<https://papers.nips.cc/paper/6335-deep-alternative-neural-network-exploring-contexts-as-early-as-possible-for-action-recognition>  
AUTHORS: Jinzhuo Wang, Wenmin Wang, xiongtao Chen, Ronggang Wang, Wen Gao  
HIGHLIGHT: Besides, we present an adaptive method to determine the temporal size for network input based on optical flow energy, and develop a volumetric pyramid pooling layer to deal with input clips of arbitrary sizes.
- 92, TITLE: Dual Learning for Machine Translation  
<https://papers.nips.cc/paper/6469-dual-learning-for-machine-translation>  
AUTHORS: Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, Wei-Ying Ma  
HIGHLIGHT: To tackle this training data bottleneck, we develop a dual-learning mechanism, which can enable an NMT system to automatically learn from unlabeled data through a dual-learning game.
- 93, TITLE: Dialog-based Language Learning  
<https://papers.nips.cc/paper/6264-dialog-based-language-learning>  
AUTHORS: Jason E. Weston  
HIGHLIGHT: A long-term goal of machine learning research is to build an intelligent dialog agent.
- 94, TITLE: Joint Line Segmentation and Transcription for End-to-End Handwritten Paragraph Recognition  
<https://papers.nips.cc/paper/6257-joint-line-segmentation-and-transcription-for-end-to-end-handwritten-paragraph-recognition>  
AUTHORS: Theodore Bluche  
HIGHLIGHT: In this paper, we propose a modification of the popular and efficient Multi-Dimensional Long Short-Term Memory Recurrent Neural Networks (MDLSTM-RNNs) to enable end-to-end processing of handwritten paragraphs.
- 95, TITLE: Temporal Regularized Matrix Factorization for High-dimensional Time Series Prediction  
<https://papers.nips.cc/paper/6160-temporal-regularized-matrix-factorization-for-high-dimensional-time-series-prediction>  
AUTHORS: Hsiang-Fu Yu, Nikhil Rao, Inderjit S. Dhillon  
HIGHLIGHT: In this paper, we present a temporal regularized matrix factorization (TRMF) framework which supports data-driven temporal learning and forecasting.
- 96, TITLE: Active Nearest-Neighbor Learning in Metric Spaces  
<https://papers.nips.cc/paper/6100-active-nearest-neighbor-learning-in-metric-spaces>  
AUTHORS: Aryeh Kontorovich, Sivan Sabato, Ruth Umer  
HIGHLIGHT: We propose a pool-based non-parametric active learning algorithm for general metric spaces, called MArgin Regularized Metric Active Nearest Neighbor (MARMANN), which outputs a nearest-neighbor classifier.
- 97, TITLE: Proximal Deep Structured Models  
<https://papers.nips.cc/paper/6074-proximal-deep-structured-models>  
AUTHORS: Shenlong Wang, Sanja Fidler, Raquel Urtasun  
HIGHLIGHT: In this paper, we propose a powerful deep structured model that is able to learn complex non-linear functions which encode the dependencies between continuous output variables.
- 98, TITLE: Faster Projection-free Convex Optimization over the Spectrahedron  
<https://papers.nips.cc/paper/6397-faster-projection-free-convex-optimization-over-the-spectrahedron>  
AUTHORS: Dan Garber, Dan Garber  
HIGHLIGHT: In this work we present a modification of the CG method tailored for the spectrahedron.
- 99, TITLE: Bayesian Optimization with a Finite Budget: An Approximate Dynamic Programming Approach  
<https://papers.nips.cc/paper/6188-bayesian-optimization-with-a-finite-budget-an-approximate-dynamic-programming-approach>  
AUTHORS: Remi Lam, Karen Willcox, David H. Wolpert  
HIGHLIGHT: We consider the problem of optimizing an expensive objective function when a finite budget of total evaluations is prescribed.
- 100, TITLE: SoundNet: Learning Sound Representations from Unlabeled Video

<https://papers.nips.cc/paper/6146-soundnet-learning-sound-representations-from-unlabeled-video>

AUTHORS: Yusuf Aytar, Carl Vondrick, Antonio Torralba

HIGHLIGHT: We propose a student-teacher training procedure which transfers discriminative visual knowledge from well established visual recognition models into the sound modality using unlabeled video as a bridge.

101, TITLE: Weight Normalization: A Simple Reparameterization to Accelerate Training of Deep Neural Networks

<https://papers.nips.cc/paper/6114-weight-normalization-a-simple-reparameterization-to-accelerate-training-of-deep-neural-networks>

AUTHORS: Tim Salimans, Durk P. Kingma

HIGHLIGHT: We present weight normalization: a reparameterization of the weight vectors in a neural network that decouples the length of those weight vectors from their direction.

102, TITLE: Efficient Second Order Online Learning by Sketching

<https://papers.nips.cc/paper/6207-efficient-second-order-online-learning-by-sketching>

AUTHORS: Haipeng Luo, Alekh Agarwal, Nicolò Cesa-Bianchi, John Langford

HIGHLIGHT: We propose Sketched Online Newton (SON), an online second order learning algorithm that enjoys substantially improved regret guarantees for ill-conditioned data.

103, TITLE: Dynamic Mode Decomposition with Reproducing Kernels for Koopman Spectral Analysis

<https://papers.nips.cc/paper/6583-dynamic-mode-decomposition-with-reproducing-kernels-for-koopman-spectral-analysis>

AUTHORS: Yoshinobu Kawahara

HIGHLIGHT: In this paper, we consider a spectral analysis of the Koopman operator in a reproducing kernel Hilbert space (RKHS).

104, TITLE: Distributed Flexible Nonlinear Tensor Factorization

<https://papers.nips.cc/paper/6585-distributed-flexible-nonlinear-tensor-factorization>

AUTHORS: Shandian Zhe, Kai Zhang, Pengyuan Wang, Kuang-chih Lee, Zenglin Xu, Yuan Qi, Zoubin Ghahramani

HIGHLIGHT: Therefore, we propose a distributed, flexible nonlinear tensor factorization model, which avoids the expensive computations and structural restrictions of the Kronecker-product in the existing TGP formulations, allowing an arbitrary subset of tensor entries to be selected for training.

105, TITLE: The Robustness of Estimator Composition

<https://papers.nips.cc/paper/6056-the-robustness-of-estimator-composition>

AUTHORS: Pingfan Tang, Jeff M. Phillips

HIGHLIGHT: We formalize notions of robustness for composite estimators via the notion of a breakdown point.

106, TITLE: Efficient and Robust Spiking Neural Circuit for Navigation Inspired by Echolocating Bats

<https://papers.nips.cc/paper/6558-efficient-and-robust-spiking-neural-circuit-for-navigation-inspired-by-echolocating-bats>

AUTHORS: Pulkit Tandon, Yash H. Malviya, Bipin Rajendran

HIGHLIGHT: We demonstrate a spiking neural circuit for azimuth angle detection inspired by the echolocation circuits of the Horseshoe bat *Rhinolophus ferrumequinum* and utilize it to devise a model for navigation and target tracking, capturing several key aspects of information transmission in biology.

107, TITLE: PerforatedCNNs: Acceleration through Elimination of Redundant Convolutions

<https://papers.nips.cc/paper/6463-perforatedcnns-acceleration-through-elimination-of-redundant-convolutions>

AUTHORS: Mikhail Figurnov, Aizhan Ibraimova, Dmitry P. Vetrov, Pushmeet Kohli

HIGHLIGHT: We propose a novel approach to reduce the computational cost of evaluation of convolutional neural networks, a factor that has hindered their deployment in low-power devices such as mobile phones.

108, TITLE: Differential Privacy without Sensitivity

<https://papers.nips.cc/paper/6050-differential-privacy-without-sensitivity>

AUTHORS: Kentaro Minami, Hltomi Arai, Issei Sato, Hiroshi Nakagawa

HIGHLIGHT: In this paper, we focus on  $(\epsilon, \delta)$ -differential privacy of Gibbs posteriors with convex and Lipschitz loss functions.

109, TITLE: Optimal Cluster Recovery in the Labeled Stochastic Block Model

<https://papers.nips.cc/paper/6196-optimal-cluster-recovery-in-the-labeled-stochastic-block-model>

AUTHORS: Se-Young Yun, Alexandre Proutiere

HIGHLIGHT: We consider the problem of community detection or clustering in the labeled Stochastic Block Model (LSBM) with a finite number  $K$  of clusters of sizes linearly growing with the global population of items  $n$ .

- 110, TITLE: LazySVD: Even Faster SVD Decomposition Yet Without Agonizing Pain  
<https://papers.nips.cc/paper/6507-lazysvd-even-faster-svd-decomposition-yet-without-agonizing-pain>  
AUTHORS: Zeyuan Allen-Zhu, Yuanzhi Li  
HIGHLIGHT: In this paper, we put forward a new and simple LazySVD framework to improve the above breakthroughs.
- 111, TITLE: An algorithm for L1 nearest neighbor search via monotonic embedding  
<https://papers.nips.cc/paper/6227-an-algorithm-for-l1-nearest-neighbor-search-via-monotonic-embedding>  
AUTHORS: Xinan Wang, Sanjoy Dasgupta  
HIGHLIGHT: Here we develop an approach for L1 distance that begins with an explicit and exact embedding of the points into L2.
- 112, TITLE: Gaussian Process Bandit Optimisation with Multi-fidelity Evaluations  
<https://papers.nips.cc/paper/6118-gaussian-process-bandit-optimisation-with-multi-fidelity-evaluations>  
AUTHORS: Kirthevasan Kandasamy, Gautam Dasarathy, Junier B. Oliva, Jeff Schneider, Barnabas Poczos  
HIGHLIGHT: We formalise this task as a **multi-fidelity** bandit problem where the target function and its approximations are sampled from a Gaussian process.
- 113, TITLE: Linear-Memory and Decomposition-Invariant Linearly Convergent Conditional Gradient Algorithm for Structured Polytopes  
<https://papers.nips.cc/paper/6115-linear-memory-and-decomposition-invariant-linearly-convergent-conditional-gradient-algorithm-for-structured-polytopes>  
AUTHORS: Dan Garber, Dan Garber, Ofer Meshi  
HIGHLIGHT: Recently, several works have shown that natural modifications of the classical conditional gradient method (aka Frank-Wolfe algorithm) for constrained convex optimization, provably converge with a linear rate when the feasible set is a polytope, and the objective is smooth and strongly-convex.
- 114, TITLE: Efficient Nonparametric Smoothness Estimation  
<https://papers.nips.cc/paper/6369-efficient-nonparametric-smoothness-estimation>  
AUTHORS: Shashank Singh, Simon S. Du, Barnabas Poczos  
HIGHLIGHT: We propose and analyze a family of estimators for Sobolev quantities of unknown probability density functions.
- 115, TITLE: A Theoretically Grounded Application of Dropout in Recurrent Neural Networks  
<https://papers.nips.cc/paper/6241-a-theoretically-grounded-application-of-dropout-in-recurrent-neural-networks>  
AUTHORS: Yarin Gal, Zoubin Ghahramani  
HIGHLIGHT: We apply this new variational inference based dropout technique in LSTM and GRU models, assessing it on language modelling and sentiment analysis tasks.
- 116, TITLE: Fast e-free Inference of Simulation Models with Bayesian Conditional Density Estimation  
<https://papers.nips.cc/paper/6084-fast-free-inference-of-simulation-models-with-bayesian-conditional-density-estimation>  
AUTHORS: George Papamakarios, Iain Murray  
HIGHLIGHT: We propose a new approach to likelihood-free inference based on Bayesian conditional density estimation.
- 117, TITLE: Direct Feedback Alignment Provides Learning in Deep Neural Networks  
<https://papers.nips.cc/paper/6441-direct-feedback-alignment-provides-learning-in-deep-neural-networks>  
AUTHORS: Arild N?kland  
HIGHLIGHT: In this work, the feedback alignment principle is used for training hidden layers more independently from the rest of the network, and from a zero initial condition.
- 118, TITLE: Safe and Efficient Off-Policy Reinforcement Learning  
<https://papers.nips.cc/paper/6538-safe-and-efficient-off-policy-reinforcement-learning>  
AUTHORS: Remi Munos, Tom Stepleton, Anna Harutyunyan, Marc Bellemare  
HIGHLIGHT: In this work, we take a fresh look at some old and new algorithms for off-policy, return-based reinforcement learning.
- 119, TITLE: A Multi-Batch L-BFGS Method for Machine Learning  
<https://papers.nips.cc/paper/6145-a-multi-batch-l-bfgs-method-for-machine-learning>  
AUTHORS: Albert S. Berahas, Jorge Nocedal, Martin Takac  
HIGHLIGHT: In this paper, we focus instead on batch methods that use a sizeable fraction of the training set at each iteration to facilitate parallelism, and that employ second-order information.

- 120, TITLE: Semiparametric Differential Graph Models  
<https://papers.nips.cc/paper/6529-semiparametric-differential-graph-models>  
AUTHORS: Pan Xu, Quanquan Gu  
HIGHLIGHT: We propose a novel graphical model, namely Latent Differential Graph Model, where the networks under two different conditions are represented by two semiparametric elliptical distributions respectively, and the variation of these two networks (i.e., differential graph) is characterized by the difference between their latent precision matrices.
- 121, TITLE: Rényi Divergence Variational Inference  
<https://papers.nips.cc/paper/6208-renyi-divergence-variational-inference>  
AUTHORS: Yingzhen Li, Richard E. Turner  
HIGHLIGHT: This paper introduces the variational Rényi bound (VR) that extends traditional variational inference to Rényi's alpha-divergences.
- 122, TITLE: Doubly Convolutional Neural Networks  
<https://papers.nips.cc/paper/6340-doubly-convolutional-neural-networks>  
AUTHORS: Shuangfei Zhai, Yu Cheng, Zhongfei (Mark) Zhang, Weining Lu  
HIGHLIGHT: In this paper, we propose doubly convolutional neural networks (DCNNs), which significantly improve the performance of CNNs by further exploring this idea.
- 123, TITLE: Density Estimation via Discrepancy Based Adaptive Sequential Partition  
<https://papers.nips.cc/paper/6217-density-estimation-via-discrepancy-based-adaptive-sequential-partition>  
AUTHORS: Dangna Li, Kun Yang, Wing Hung Wong  
HIGHLIGHT: Given  $n$  observations from an unknown continuous distribution defined on some domain  $\Omega$ , we propose a nonparametric method to learn a piecewise constant function to approximate the underlying probability density function.
- 124, TITLE: How Deep is the Feature Analysis underlying Rapid Visual Categorization?  
<https://papers.nips.cc/paper/6218-how-deep-is-the-feature-analysis-underlying-rapid-visual-categorization>  
AUTHORS: Sven Eberhardt, Jonah G. Cader, Thomas Serre  
HIGHLIGHT: We have conducted a large-scale psychophysics study to assess the correlation between computational models and human behavioral responses on a rapid animal vs. non-animal categorization task.
- 125, TITLE: VIME: Variational Information Maximizing Exploration  
<https://papers.nips.cc/paper/6591-vime-variational-information-maximizing-exploration>  
AUTHORS: Rein Houthooft, Xi Chen, Xi Chen, Yan Duan, John Schulman, Filip De Turck, Pieter Abbeel  
HIGHLIGHT: This paper introduces Variational Information Maximizing Exploration (VIME), an exploration strategy based on maximization of information gain about the agent's belief of environment dynamics.
- 126, TITLE: Generalized Correspondence-LDA Models (GC-LDA) for Identifying Functional Regions in the Brain  
<https://papers.nips.cc/paper/6274-generalized-correspondence-lda-models-gc-lda-for-identifying-functional-regions-in-the-brain>  
AUTHORS: Timothy Rubin, Oluwasanmi O. Koyejo, Michael N. Jones, Tal Yarkoni  
HIGHLIGHT: This paper presents Generalized Correspondence-LDA (GC-LDA), a generalization of the Correspondence-LDA model that allows for variable spatial representations to be associated with topics, and increased flexibility in terms of the strength of the correspondence between data types induced by the model.
- 127, TITLE: Solving Marginal MAP Problems with NP Oracles and Parity Constraints  
<https://papers.nips.cc/paper/6462-solving-marginal-map-problems-with-np-oracles-and-parity-constraints>  
AUTHORS: Yexiang Xue, Zhiyuan Li, Stefano Ermon, Carla P. Gomes, Bart Selman  
HIGHLIGHT: We propose XOR\_MMAP, a novel approach to solve the Marginal MAP problem, which represents the intractable counting subproblem with queries to NP oracles, subject to additional parity constraints.
- 128, TITLE: Multi-view Anomaly Detection via Robust Probabilistic Latent Variable Models  
<https://papers.nips.cc/paper/6456-multi-view-anomaly-detection-via-robust-probabilistic-latent-variable-models>  
AUTHORS: Tomoharu Iwata, Makoto Yamada  
HIGHLIGHT: We propose probabilistic latent variable models for multi-view anomaly detection, which is the task of finding instances that have inconsistent views given multi-view data.
- 129, TITLE: Proximal Stochastic Methods for Nonsmooth Nonconvex Finite-Sum Optimization  
<https://papers.nips.cc/paper/6116-proximal-stochastic-methods-for-nonsmooth-nonconvex-finite-sum-optimization>  
AUTHORS: Sashank J. Reddi, Suvrit Sra, Barnabas Poczos, Alexander J. Smola  
HIGHLIGHT: We analyze stochastic algorithms for optimizing nonconvex, nonsmooth finite-sum problems, where the nonsmooth part is convex.

- 130, TITLE: Variance Reduction in Stochastic Gradient Langevin Dynamics  
<https://papers.nips.cc/paper/6293-variance-reduction-in-stochastic-gradient-langevin-dynamics>  
AUTHORS: Kumar Avinava Dubey, Sashank J. Reddi, Sinead A. Williamson, Barnabas Poczos, Alexander J. Smola, Eric P. Xing  
HIGHLIGHT: In this paper, we present techniques for reducing variance in stochastic gradient Langevin dynamics, yielding novel stochastic Monte Carlo methods that improve performance by reducing the variance in the stochastic gradient.
- 131, TITLE: Regularization With Stochastic Transformations and Perturbations for Deep Semi-Supervised Learning  
<https://papers.nips.cc/paper/6333-regularization-with-stochastic-transformations-and-perturbations-for-deep-semi-supervised-learning>  
AUTHORS: Mehdi Sajjadi, Mehran Javanmardi, Tolga Tasdizen  
HIGHLIGHT: In this paper, we consider the problem of semi-supervised learning with convolutional neural networks.
- 132, TITLE: Dense Associative Memory for Pattern Recognition  
<https://papers.nips.cc/paper/6121-dense-associative-memory-for-pattern-recognition>  
AUTHORS: Dmitry Krotov, John J. Hopfield  
HIGHLIGHT: We propose a simple duality between this dense associative memory and neural networks commonly used in deep learning.
- 133, TITLE: Causal Bandits: Learning Good Interventions via Causal Inference  
<https://papers.nips.cc/paper/6195-causal-bandits-learning-good-interventions-via-causal-inference>  
AUTHORS: Finnian Lattimore, Tor Lattimore, Mark D. Reid  
HIGHLIGHT: We propose a new algorithm that exploits the causal feedback and prove a bound on its simple regret that is strictly better (in all quantities) than algorithms that do not use the additional causal information.
- 134, TITLE: Refined Lower Bounds for Adversarial Bandits  
<https://papers.nips.cc/paper/6277-refined-lower-bounds-for-adversarial-bandits>  
AUTHORS: S?bastien Gerchinovitz, Tor Lattimore  
HIGHLIGHT: We provide new lower bounds on the regret that must be suffered by adversarial bandit algorithms.
- 135, TITLE: Theoretical Comparisons of Positive-Unlabeled Learning against Positive-Negative Learning  
<https://papers.nips.cc/paper/6354-theoretical-comparisons-of-positive-unlabeled-learning-against-positive-negative-learning>  
AUTHORS: Gang Niu, Marthinus Christoffel du Plessis, Tomoya Sakai, Yao Ma, Masashi Sugiyama  
HIGHLIGHT: In this paper, we theoretically compare PU (and NU) learning against PN learning based on the upper bounds on estimation errors.
- 136, TITLE: Homotopy Smoothing for Non-Smooth Problems with Lower Complexity than  $O(1/\epsilon)$   
<https://papers.nips.cc/paper/6407-homotopy-smoothing-for-non-smooth-problems-with-lower-complexity-than-o-1-epsilon>  
AUTHORS: Yi Xu, Yan Yan, Qihang Lin, Tianbao Yang  
HIGHLIGHT: In this paper, we develop a novel  $\{bf ho\}$   $\{bf p\}$   $\{bf s\}$  smoothing (HOPS) algorithm for solving a family of non-smooth problems that is composed of a non-smooth term with an explicit max-structure and a smooth term or a simple non-smooth term whose proximal mapping is easy to compute.
- 137, TITLE: Finite-Sample Analysis of Fixed-k Nearest Neighbor Density Functional Estimators  
<https://papers.nips.cc/paper/6123-finite-sample-analysis-of-fixed-k-nearest-neighbor-density-functional-estimators>  
AUTHORS: Shashank Singh, Barnabas Poczos  
HIGHLIGHT: We provide finite-sample analysis of a general framework for using k-nearest neighbor statistics to estimate functionals of a nonparametric continuous probability density, including entropies and divergences.
- 138, TITLE: A state-space model of cross-region dynamic connectivity in MEG/EEG  
<https://papers.nips.cc/paper/6593-a-state-space-model-of-cross-region-dynamic-connectivity-in-megeeg>  
AUTHORS: Ying Yang, Elissa Aminoff, Michael Tarr, Kass E. Robert  
HIGHLIGHT: In this work, we propose a one-step state-space model to improve estimation of dynamic connectivity.
- 139, TITLE: What Makes Objects Similar: A Unified Multi-Metric Learning Approach  
<https://papers.nips.cc/paper/6192-what-makes-objects-similar-a-unified-multi-metric-learning-approach>  
AUTHORS: Han-Jia Ye, De-Chuan Zhan, Xue-Min Si, Yuan Jiang, Zhi-Hua Zhou  
HIGHLIGHT: We propose a Unified Multi-Metric Learning (UM2L) framework to exploit multiple types of metrics.

- 140, TITLE: Adaptive Maximization of Pointwise Submodular Functions With Budget Constraint  
<https://papers.nips.cc/paper/6526-adaptive-maximization-of-pointwise-submodular-functions-with-budget-constraint>  
AUTHORS: Nguyen Cuong, Huan Xu  
HIGHLIGHT: We study the worst-case adaptive optimization problem with budget constraint that is useful for modeling various practical applications in artificial intelligence and machine learning.
- 141, TITLE: Dueling Bandits: Beyond Condorcet Winners to General Tournament Solutions  
<https://papers.nips.cc/paper/6337-dueling-bandits-beyond-condorcet-winners-to-general-tournament-solutions>  
AUTHORS: Siddhartha Y. Ramamohan, Arun Rajkumar, Shivani Agarwal, Shivani Agarwal  
HIGHLIGHT: In this work, we consider a broad notion of winners defined by tournament solutions in social choice theory, which include the Copeland set as a special case but also include several other notions of winners such as the top cycle, uncovered set, and Banks set, and which, like the Copeland set, always exist.
- 142, TITLE: Local Similarity-Aware Deep Feature Embedding  
<https://papers.nips.cc/paper/6368-local-similarity-aware-deep-feature-embedding>  
AUTHORS: Chen Huang, Chen Change Loy, Xiaou Tang  
HIGHLIGHT: In this paper, we introduce a Position-Dependent Deep Metric (PDDM) unit, which is capable of learning a similarity metric adaptive to local feature structure.
- 143, TITLE: A Communication-Efficient Parallel Algorithm for Decision Tree  
<https://papers.nips.cc/paper/6381-a-communication-efficient-parallel-algorithm-for-decision-tree>  
AUTHORS: Qi Meng, Guolin Ke, Taifeng Wang, Wei Chen, Qiwei Ye, Zhi-Ming Ma, Tie-Yan Liu  
HIGHLIGHT: In this paper, we propose a new algorithm, called `\emph{Parallel Voting Decision Tree (PV-Tree)}`, to tackle this challenge.
- 144, TITLE: Convex Two-Layer Modeling with Latent Structure  
<https://papers.nips.cc/paper/6314-convex-two-layer-modeling-with-latent-structure>  
AUTHORS: Vignesh Ganapathiraman, Xinhua Zhang, Yaoliang Yu, Junfeng Wen  
HIGHLIGHT: In this paper, we develop a convex relaxation of two-layer conditional model which captures latent structure and estimates model parameters, jointly and optimally.
- 145, TITLE: Sampling for Bayesian Program Learning  
<https://papers.nips.cc/paper/6082-sampling-for-bayesian-program-learning>  
AUTHORS: Kevin Ellis, Armando Solar-Lezama, Josh Tenenbaum  
HIGHLIGHT: Within this setting, we propose an algorithm that uses a symbolic solver to efficiently sample programs. Towards learning programs from data, we introduce the problem of sampling programs from posterior distributions conditioned on that data.
- 146, TITLE: Learning Kernels with Random Features  
<https://papers.nips.cc/paper/6180-learning-kernels-with-random-features>  
AUTHORS: Aman Sinha, John C. Duchi  
HIGHLIGHT: Specifically, we present an efficient optimization problem that learns a kernel in a supervised manner.
- 147, TITLE: Optimal Tagging with Markov Chain Optimization  
<https://papers.nips.cc/paper/6041-optimal-tagging-with-markov-chain-optimization>  
AUTHORS: Nir Rosenfeld, Amir Globerson  
HIGHLIGHT: In this paper we introduce the problem of optimal tagging, where the task is to choose a subset of tags for a new item such that the probability of browsing users reaching that item is maximized.
- 148, TITLE: Crowdsourced Clustering: Querying Edges vs Triangles  
<https://papers.nips.cc/paper/6499-crowdsourced-clustering-querying-edges-vs-triangles>  
AUTHORS: Ramya Korlakai Vinayak, Babak Hassibi  
HIGHLIGHT: We consider the task of clustering items using answers from non-expert crowd workers.
- 149, TITLE: Mixed vine copulas as joint models of spike counts and local field potentials  
<https://papers.nips.cc/paper/6069-mixed-vine-copulas-as-joint-models-of-spike-counts-and-local-field-potentials>  
AUTHORS: Arno Onken, Stefano Panzeri  
HIGHLIGHT: Here we introduce such techniques in a framework based on vine copulas with mixed margins to construct multivariate stochastic models.

- 150, TITLE: Achieving the KS threshold in the general stochastic block model with linearized acyclic belief propagation  
<https://papers.nips.cc/paper/6365-achieving-the-ks-threshold-in-the-general-stochastic-block-model-with-linearized-acyclic-belief-propagation>  
AUTHORS: Emmanuel Abbe, Colin Sandon  
HIGHLIGHT: The stochastic block model (SBM) has long been studied in machine learning and network science as a canonical model for clustering and community detection.
- 151, TITLE: Adaptive Concentration Inequalities for Sequential Decision Problems  
<https://papers.nips.cc/paper/6493-adaptive-concentration-inequalities-for-sequential-decision-problems>  
AUTHORS: Shengjia Zhao, Enze Zhou, Ashish Sabharwal, Stefano Ermon  
HIGHLIGHT: We introduce Hoeffding-like concentration inequalities that hold for a random, adaptively chosen number of samples.
- 152, TITLE: Nested Mini-Batch K-Means  
<https://papers.nips.cc/paper/6481-nested-mini-batch-k-means>  
AUTHORS: James Newling, François Fleuret  
HIGHLIGHT: A new algorithm is proposed which accelerates the mini-batch k-means algorithm of Sculley (2010) by using the distance bounding approach of Elkan (2003).
- 153, TITLE: Deep Learning Models of the Retinal Response to Natural Scenes  
<https://papers.nips.cc/paper/6388-deep-learning-models-of-the-retinal-response-to-natural-scenes>  
AUTHORS: Lane McIntosh, Niru Maheswaranathan, Aran Nayebi, Surya Ganguli, Stephen Baccus  
HIGHLIGHT: Here we demonstrate that deep convolutional neural networks (CNNs) capture retinal responses to natural scenes nearly to within the variability of a cell's response, and are markedly more accurate than linear-nonlinear (LN) models and Generalized Linear Models (GLMs).
- 154, TITLE: Preference Completion from Partial Rankings  
<https://papers.nips.cc/paper/6272-preference-completion-from-partial-rankings>  
AUTHORS: Suriya Gunasekar, Oluwasanmi O. Koyejo, Joydeep Ghosh  
HIGHLIGHT: We propose a novel and efficient algorithm for the collaborative preference completion problem, which involves jointly estimating individualized rankings for a set of entities over a shared set of items, based on a limited number of observed affinity values.
- 155, TITLE: Dynamic Network Surgery for Efficient DNNs  
<https://papers.nips.cc/paper/6165-dynamic-network-surgery-for-efficient-dnns>  
AUTHORS: Yiwen Guo, Anbang Yao, Yurong Chen  
HIGHLIGHT: In this paper, we propose a novel network compression method called dynamic network surgery, which can remarkably reduce the network complexity by making on-the-fly connection pruning.
- 156, TITLE: Learning a Metric Embedding for Face Recognition using the Multibatch Method  
<https://papers.nips.cc/paper/6524-learning-a-metric-embedding-for-face-recognition-using-the-multibatch-method>  
AUTHORS: Oren Tadmor, Tal Rosenwein, Shai Shalev-Shwartz, Yonatan Wexler, Amnon Shashua  
HIGHLIGHT: We prove that the variance of the Multibatch estimator is bounded by  $\mathcal{O}(1/k^2)$ , under some mild conditions.
- 157, TITLE: A Pseudo-Bayesian Algorithm for Robust PCA  
<https://papers.nips.cc/paper/6435-a-pseudo-bayesian-algorithm-for-robust-pca>  
AUTHORS: Tae-Hyun Oh, Yasuyuki Matsushita, In Kweon, David Wipf  
HIGHLIGHT: In this paper, we propose a novel pseudo-Bayesian algorithm that explicitly compensates for design weaknesses in many existing non-convex approaches leading to state-of-the-art performance with a sound analytical foundation.
- 158, TITLE: End-to-End Kernel Learning with Supervised Convolutional Kernel Networks  
<https://papers.nips.cc/paper/6184-end-to-end-kernel-learning-with-supervised-convolutional-kernel-networks>  
AUTHORS: Julien Mairal  
HIGHLIGHT: In this paper, we introduce a new image representation based on a multilayer kernel machine.
- 159, TITLE: Stochastic Variance Reduction Methods for Saddle-Point Problems  
<https://papers.nips.cc/paper/6471-stochastic-variance-reduction-methods-for-saddle-point-problems>  
AUTHORS: Balamurugan Palaniappan, Francis Bach  
HIGHLIGHT: We consider convex-concave saddle-point problems where the objective functions may be split in many components, and extend recent stochastic variance reduction methods (such as SVRG or SAGA) to provide the first large-scale linearly convergent algorithms for this class of problems which are common in machine learning.

160, TITLE: Flexible Models for Microclustering with Application to Entity Resolution  
<https://papers.nips.cc/paper/6334-flexible-models-for-microclustering-with-application-to-entity-resolution>  
AUTHORS: Brenda Betancourt, Giacomo Zanella, Jeffrey W. Miller, Hanna Wallach, Abbas Zaidi, Rebecca C. Steorts  
HIGHLIGHT: We compare models within this class to two commonly used clustering models using four entity-resolution data sets.

161, TITLE: Catching heuristics are optimal control policies  
<https://papers.nips.cc/paper/6548-catching-heuristics-are-optimal-control-policies>  
AUTHORS: Boris Belousov, Gerhard Neumann, Constantin A. Rothkopf, Jan R. Peters  
HIGHLIGHT: In this paper, we show that interception strategies appearing to be heuristics can be understood as computational solutions to the optimal control problem faced by a ball-catching agent acting under uncertainty.

162, TITLE: Bayesian optimization under mixed constraints with a slack-variable augmented Lagrangian  
<https://papers.nips.cc/paper/6439-bayesian-optimization-under-mixed-constraints-with-a-slack-variable-augmented-lagrangian>  
AUTHORS: Victor Picheny, Robert B. Gramacy, Stefan Wild, Sebastien Le Digabel  
HIGHLIGHT: Here we introduce an alternative slack variable AL, and show that in this formulation the EI may be evaluated with library routines.

163, TITLE: Adaptive Neural Compilation  
<https://papers.nips.cc/paper/6411-adaptive-neural-compilation>  
AUTHORS: Rudy R. Bunel, Alban Desmaison, Pawan K. Mudigonda, Pushmeet Kohli, Philip Torr  
HIGHLIGHT: This paper proposes an adaptive neural-compilation framework to address the problem of learning efficient program.

164, TITLE: Synthesis of MCMC and Belief Propagation  
<https://papers.nips.cc/paper/6318-synthesis-of-mcmc-and-belief-propagation>  
AUTHORS: Sung-Soo Ahn, Michael Chertkov, Jinwoo Shin  
HIGHLIGHT: In this paper, we introduce MCMC algorithms correcting the approximation error of BP, i.e., we provide a way to compensate for BP errors via a consecutive BP-aware MCMC.

165, TITLE: Learning Treewidth-Bounded Bayesian Networks with Thousands of Variables  
<https://papers.nips.cc/paper/6232-learning-treewidth-bounded-bayesian-networks-with-thousands-of-variables>  
AUTHORS: Mauro Scanagatta, Giorgio Corani, Cassio P. de Campos, Marco Zaffalon  
HIGHLIGHT: We present a method for learning treewidth-bounded Bayesian networks from data sets containing thousands of variables.

166, TITLE: Unifying Count-Based Exploration and Intrinsic Motivation  
<https://papers.nips.cc/paper/6383-unifying-count-based-exploration-and-intrinsic-motivation>  
AUTHORS: Marc Bellemare, Sriram Srinivasan, Georg Ostrovski, Tom Schaul, David Saxton, Remi Munos  
HIGHLIGHT: Drawing inspiration from the intrinsic motivation literature, we use density models to measure uncertainty, and propose a novel algorithm for deriving a pseudo-count from an arbitrary density model.

167, TITLE: Large Margin Discriminant Dimensionality Reduction in Prediction Space  
<https://papers.nips.cc/paper/6458-large-margin-discriminant-dimensionality-reduction-in-prediction-space>  
AUTHORS: Mohammad Saberian, Jose Costa Pereira, Can Xu, Jian Yang, Nuno Vasconcelos  
HIGHLIGHT: In this paper we establish a duality between boosting and SVM, and use this to derive a novel discriminant dimensionality reduction algorithm.

168, TITLE: Stochastic Structured Prediction under Bandit Feedback  
<https://papers.nips.cc/paper/6134-stochastic-structured-prediction-under-bandit-feedback>  
AUTHORS: Artem Sokolov, Julia Kreutzer, Stefan Riezler, Christopher Lo  
HIGHLIGHT: We present applications of this learning scenario to convex and non-convex objectives for structured prediction and analyze them as stochastic first-order methods.

169, TITLE: Simple and Efficient Weighted Minwise Hashing  
<https://papers.nips.cc/paper/6472-simple-and-efficient-weighted-minwise-hashing>  
AUTHORS: Anshumali Shrivastava  
HIGHLIGHT: We propose a simple rejection type sampling scheme based on a carefully designed red-green map, where we show that the number of rejected sample has exactly the same distribution as weighted minwise sampling.

170, TITLE: Truncated Variance Reduction: A Unified Approach to Bayesian Optimization and Level-Set Estimation  
<https://papers.nips.cc/paper/6080-truncated-variance-reduction-a-unified-approach-to-bayesian-optimization-and-level-set-estimation>  
AUTHORS: Ilija Bogunovic, Jonathan Scarlett, Andreas Krause, Volkan Cevher  
HIGHLIGHT: We present a new algorithm, truncated variance reduction (TruVaR), that treats Bayesian optimization (BO) and level-set estimation (LSE) with Gaussian processes in a unified fashion.

171, TITLE: Structured Sparse Regression via Greedy Hard Thresholding  
<https://papers.nips.cc/paper/6425-structured-sparse-regression-via-greedy-hard-thresholding>  
AUTHORS: Prateek Jain, Nikhil Rao, Inderjit S. Dhillon  
HIGHLIGHT: In this paper, we show that such NP-hard projections can not only be avoided by appealing to submodular optimization, but such methods come with strong theoretical guarantees even in the presence of poorly conditioned data (i.e. say when two features have correlation  $\geq 0.99$ ), which existing analyses cannot handle.

172, TITLE: Understanding Probabilistic Sparse Gaussian Process Approximations  
<https://papers.nips.cc/paper/6477-understanding-probabilistic-sparse-gaussian-process-approximations>  
AUTHORS: Matthias Bauer, Mark van der Wilk, Carl Edward Rasmussen  
HIGHLIGHT: We thoroughly investigate the two methods for regression both analytically and through illustrative examples, and draw conclusions to guide practical application.

173, TITLE: SEBOOST - Boosting Stochastic Learning Using Subspace Optimization Techniques  
<https://papers.nips.cc/paper/6109-seboost-boosting-stochastic-learning-using-subspace-optimization-techniques>  
AUTHORS: Elad Richardson, Rom Herskovitz, Boris Ginsburg, Michael Zibulevsky  
HIGHLIGHT: We present SEBOOST, a technique for boosting the performance of existing stochastic optimization methods.

174, TITLE: Generating Long-term Trajectories Using Deep Hierarchical Networks  
<https://papers.nips.cc/paper/6520-generating-long-term-trajectories-using-deep-hierarchical-networks>  
AUTHORS: Stephan Zheng, Yisong Yue, Jennifer Hobbs  
HIGHLIGHT: We instead propose a hierarchical policy class that automatically reasons about both long-term and short-term goals, which we instantiate as a hierarchical neural network.

175, TITLE: Learning Tree Structured Potential Games  
<https://papers.nips.cc/paper/6152-learning-tree-structured-potential-games>  
AUTHORS: Vikas Garg, Tommi Jaakkola  
HIGHLIGHT: We cast the learning problem within a max margin setting and show that the problem is NP-hard even when the strategic interactions form a tree.

176, TITLE: Observational-Interventional Priors for Dose-Response Learning  
<https://papers.nips.cc/paper/6107-observational-interventional-priors-for-dose-response-learning>  
AUTHORS: Ricardo Silva  
HIGHLIGHT: In this paper, we introduce a hierarchical Gaussian process prior that constructs a distribution over the dose-response curve by learning from observational data, and reshapes the distribution with a nonparametric affine transform learned from controlled interventions.

177, TITLE: Learning from Rational Behavior: Predicting Solutions to Unknown Linear Programs  
<https://papers.nips.cc/paper/6150-learning-from-rational-behavior-predicting-solutions-to-unknown-linear-programs>  
AUTHORS: Shahin Jabbari, Ryan M. Rogers, Aaron Roth, Steven Z. Wu  
HIGHLIGHT: We define and study the problem of predicting the solution to a linear program (LP) given only partial information about its objective and constraints.

178, TITLE: Identification and Overidentification of Linear Structural Equation Models  
<https://papers.nips.cc/paper/6223-identification-and-overidentification-of-linear-structural-equation-models>  
AUTHORS: Bryant Chen  
HIGHLIGHT: In this paper, we address the problems of identifying linear structural equation models and discovering the constraints they imply.

179, TITLE: Adaptive Skills Adaptive Partitions (ASAP)  
<https://papers.nips.cc/paper/6350-adaptive-skills-adaptive-partitions-asap>  
AUTHORS: Daniel J. Mankowitz, Timothy A. Mann, Shie Mannor

**HIGHLIGHT:** We introduce the Adaptive Skills, Adaptive Partitions (ASAP) framework that (1) learns skills (i.e., temporally extended actions or options) as well as (2) where to apply them.

180, **TITLE:** Multiple-Play Bandits in the Position-Based Model  
<https://papers.nips.cc/paper/6546-multiple-play-bandits-in-the-position-based-model>

**AUTHORS:** Paul Lagr?e, Claire Vernade, Olivier Cappe

**HIGHLIGHT:** The present work proposes to exploit available information regarding the display position bias under the so-called Position-based click model (PBM).

181, **TITLE:** Optimal Black-Box Reductions Between Optimization Objectives  
<https://papers.nips.cc/paper/6364-optimal-black-box-reductions-between-optimization-objectives>

**AUTHORS:** Zeyuan Allen-Zhu, Elad Hazan

**HIGHLIGHT:** We reduce the complexity of algorithm design for machine learning by reductions: we develop reductions that take a method developed for one setting and apply it to the entire spectrum of smoothness and strong-convexity in applications.

182, **TITLE:** On Valid Optimal Assignment Kernels and Applications to Graph Classification  
<https://papers.nips.cc/paper/6166-on-valid-optimal-assignment-kernels-and-applications-to-graph-classification>

**AUTHORS:** Nils M. Kriege, Pierre-Louis Giscard, Richard Wilson

**HIGHLIGHT:** We characterize a class of base kernels used to compare parts that guarantees positive semidefinite optimal assignment kernels.

183, **TITLE:** Robustness of classifiers: from adversarial to random noise  
<https://papers.nips.cc/paper/6331-robustness-of-classifiers-from-adversarial-to-random-noise>

**AUTHORS:** Alhussein Fawzi, Seyed-Mohsen Moosavi-Dezfooli, Pascal Frossard

**HIGHLIGHT:** In this paper, we propose to study a semi-random noise regime that generalizes both the random and worst-case noise regimes.

184, **TITLE:** A Non-convex One-Pass Framework for Generalized Factorization Machine and Rank-One Matrix Sensing  
<https://papers.nips.cc/paper/6410-a-non-convex-one-pass-framework-for-generalized-factorization-machine-and-rank-one-matrix-sensing>

**AUTHORS:** Ming Lin, Jieping Ye

**HIGHLIGHT:** We develop an efficient alternating framework for learning a generalized version of Factorization Machine (gFM) on streaming data with provable guarantees.

185, **TITLE:** Exploiting the Structure: Stochastic Gradient Methods Using Raw Clusters  
<https://papers.nips.cc/paper/6403-exploiting-the-structure-stochastic-gradient-methods-using-raw-clusters>

**AUTHORS:** Zeyuan Allen-Zhu, Yang Yuan, Karthik Sridharan

**HIGHLIGHT:** In this paper we focus on one of the fundamental machine learning tasks, empirical risk minimization (ERM), and provide faster algorithms with the help from the clustering structure of the data.

186, **TITLE:** Combinatorial Multi-Armed Bandit with General Reward Functions  
<https://papers.nips.cc/paper/6511-combinatorial-multi-armed-bandit-with-general-reward-functions>

**AUTHORS:** Wei Chen, Wei Hu, Fu Li, Jian Li, Yu Liu, Pinyan Lu

**HIGHLIGHT:** In this paper, we study the stochastic combinatorial multi-armed bandit (CMAB) framework that allows a general nonlinear reward function, whose expected value may not depend only on the means of the input random variables but possibly on the entire distributions of these variables.

187, **TITLE:** Boosting with Abstention  
<https://papers.nips.cc/paper/6336-boosting-with-abstention>

**AUTHORS:** Corinna Cortes, Giulia DeSalvo, Mehryar Mohri

**HIGHLIGHT:** We present a new boosting algorithm for the key scenario of binary classification with abstention where the algorithm can abstain from predicting the label of a point, at the price of a fixed cost.

188, **TITLE:** Regret of Queueing Bandits  
<https://papers.nips.cc/paper/6370-regret-of-queueing-bandits>

**AUTHORS:** Subhashini Krishnasamy, Rajat Sen, Ramesh Johari, Sanjay Shakkottai

**HIGHLIGHT:** We study algorithms that minimize queue-regret: the (expected) difference between the queue-lengths obtained by the algorithm, and those obtained by a genie-aided matching algorithm that knows exact service rates.

189, **TITLE:** Deep Learning Games

- <https://papers.nips.cc/paper/6315-deep-learning-games>  
AUTHORS: Dale Schuurmans, Martin A. Zinkevich  
HIGHLIGHT: For convex one-layer problems, we demonstrate an equivalence between global minimizers of the training problem and Nash equilibria in a simple game.
- 190, TITLE: Globally Optimal Training of Generalized Polynomial Neural Networks with Nonlinear Spectral Methods  
<https://papers.nips.cc/paper/6238-globally-optimal-training-of-generalized-polynomial-neural-networks-with-nonlinear-spectral-methods>  
AUTHORS: Antoine Gautier, Quynh N. Nguyen, Matthias Hein  
HIGHLIGHT: In contrast we show under quite weak assumptions on the data that a particular class of feedforward neural networks can be trained globally optimal with a linear convergence rate.
- 191, TITLE: Perspective Transformer Nets: Learning Single-View 3D Object Reconstruction without 3D Supervision  
<https://papers.nips.cc/paper/6206-perspective-transformer-nets-learning-single-view-3d-object-reconstruction-without-3d-supervision>  
AUTHORS: Xinchen Yan, Jimei Yang, Ersin Yumer, Yijie Guo, Honglak Lee  
HIGHLIGHT: In this work, we investigate the task of single-view 3D object reconstruction from a learning agent's perspective.
- 192, TITLE: A Credit Assignment Compiler for Joint Prediction  
<https://papers.nips.cc/paper/6256-a-credit-assignment-compiler-for-joint-prediction>  
AUTHORS: Kai-Wei Chang, He He, Stephane Ross, Hal Daume III, John Langford  
HIGHLIGHT: In this paper, we show the search space can be defined by an arbitrary imperative program, turning learning to search into a credit assignment compiler.
- 193, TITLE: Accelerating Stochastic Composition Optimization  
<https://papers.nips.cc/paper/6438-accelerating-stochastic-composition-optimization>  
AUTHORS: Mengdi Wang, Ji Liu, Ethan Fang  
HIGHLIGHT: We propose a new stochastic first-order method, namely the accelerated stochastic compositional proximal gradient (ASC-PG) method, which updates based on queries to the sampling oracle using two different timescales.
- 194, TITLE: Reward Augmented Maximum Likelihood for Neural Structured Prediction  
<https://papers.nips.cc/paper/6547-reward-augmented-maximum-likelihood-for-neural-structured-prediction>  
AUTHORS: Mohammad Norouzi, Samy Bengio, zhifeng Chen, Navdeep Jaitly, Mike Schuster, Yonghui Wu, Dale Schuurmans  
HIGHLIGHT: This paper presents a simple and computationally efficient method that incorporates task reward into maximum likelihood training.
- 195, TITLE: Consistent Kernel Mean Estimation for Functions of Random Variables  
<https://papers.nips.cc/paper/6545-consistent-kernel-mean-estimation-for-functions-of-random-variables>  
AUTHORS: Carl-Johann Simon-Gabriel, Adam Scibior, Ilya O. Tolstikhin, Bernhard Schölkopf  
HIGHLIGHT: We provide a theoretical foundation for non-parametric estimation of functions of random variables using kernel mean embeddings.
- 196, TITLE: Towards Unifying Hamiltonian Monte Carlo and Slice Sampling  
<https://papers.nips.cc/paper/6147-towards-unifying-hamiltonian-monte-carlo-and-slice-sampling>  
AUTHORS: Yizhe Zhang, Xiangyu Wang, Changyou Chen, Ricardo Henao, Kai Fan, Lawrence Carin  
HIGHLIGHT: We provide a theoretical analysis of the mixing performance of such samplers, proving that in the limit of a single parameter, the MGS draws decorrelated samples from the desired target distribution.
- 197, TITLE: Scalable Adaptive Stochastic Optimization Using Random Projections  
<https://papers.nips.cc/paper/6054-scalable-adaptive-stochastic-optimization-using-random-projections>  
AUTHORS: Gabriel Krummenacher, Brian McWilliams, Yannic Kilcher, Joachim M. Buhmann, Nicolai Meinshausen  
HIGHLIGHT: We present Ada-LR and RadaGrad two computationally efficient approximations to full-matrix AdaGrad based on randomized dimensionality reduction.
- 198, TITLE: Variational Inference in Mixed Probabilistic Submodular Models  
<https://papers.nips.cc/paper/6225-variational-inference-in-mixed-probabilistic-submodular-models>  
AUTHORS: Josip Djolonga, Sebastian Tschiatschek, Andreas Krause  
HIGHLIGHT: We consider the problem of variational inference in probabilistic models with both log-submodular and log-supermodular higher-order potentials.

- 199, TITLE: Correlated-PCA: Principal Components Analysis when Data and Noise are Correlated  
<https://papers.nips.cc/paper/6598-correlated-pca-principal-components-analysis-when-data-and-noise-are-correlated>  
AUTHORS: Namrata Vaswani, Han Guo  
HIGHLIGHT: In this paper, we study the PCA problem in the setting where the data and noise can be correlated.
- 200, TITLE: The Multi-fidelity Multi-armed Bandit  
<https://papers.nips.cc/paper/6592-the-multi-fidelity-multi-armed-bandit>  
AUTHORS: Kirthevasan Kandasamy, Gautam Dasarathy, Barnabas Poczos, Jeff Schneider  
HIGHLIGHT: We study a variant of the classical stochastic  $K$ -armed bandit where observing the outcome of each arm is expensive, but cheap approximations to this outcome are available.
- 201, TITLE: Anchor-Free Correlated Topic Modeling: Identifiability and Algorithm  
<https://papers.nips.cc/paper/6308-anchor-free-correlated-topic-modeling-identifiability-and-algorithm>  
AUTHORS: Kejun Huang, Xiao Fu, Nikolaos D. Sidiropoulos  
HIGHLIGHT: This paper revisits topic modeling based on second-order moments, and proposes an anchor-free topic mining framework.
- 202, TITLE: Bootstrap Model Aggregation for Distributed Statistical Learning  
<https://papers.nips.cc/paper/6049-bootstrap-model-aggregation-for-distributed-statistical-learning>  
AUTHORS: JUN HAN, Qiang Liu  
HIGHLIGHT: In this work, we propose two variance reduction methods to correct the bootstrap noise, including a weighted  $M$ -estimator that is both statistically efficient and practically powerful.
- 203, TITLE: A scalable end-to-end Gaussian process adapter for irregularly sampled time series classification  
<https://papers.nips.cc/paper/6475-a-scalable-end-to-end-gaussian-process-adapter-for-irregularly-sampled-time-series-classification>  
AUTHORS: Steven Cheng-Xian Li, Benjamin M. Marlin  
HIGHLIGHT: We present a general framework for classification of sparse and irregularly-sampled time series.
- 204, TITLE: A Bandit Framework for Strategic Regression  
<https://papers.nips.cc/paper/6190-a-bandit-framework-for-strategic-regression>  
AUTHORS: Yang Liu, Yiling Chen  
HIGHLIGHT: In this work, we study a dynamic data acquisition process where data holders can contribute multiple times.
- 205, TITLE: Architectural Complexity Measures of Recurrent Neural Networks  
<https://papers.nips.cc/paper/6303-architectural-complexity-measures-of-recurrent-neural-networks>  
AUTHORS: Saizheng Zhang, Yuhuai Wu, Tong Che, Zhouhan Lin, Roland Memisevic, Ruslan R. Salakhutdinov, Yoshua Bengio  
HIGHLIGHT: In this paper, we systematically analyze the connecting architectures of recurrent neural networks (RNNs).
- 206, TITLE: Statistical Inference for Cluster Trees  
<https://papers.nips.cc/paper/6508-statistical-inference-for-cluster-trees>  
AUTHORS: Jisu KIM, Yen-Chi Chen, Sivaraman Balakrishnan, Alessandro Rinaldo, Larry Wasserman  
HIGHLIGHT: We introduce a partial ordering on cluster trees which we use to prune some of the statistically insignificant features of the empirical tree, yielding interpretable and parsimonious cluster trees.
- 207, TITLE: PAC Reinforcement Learning with Rich Observations  
<https://papers.nips.cc/paper/6575-pac-reinforcement-learning-with-rich-observations>  
AUTHORS: Akshay Krishnamurthy, Alekh Agarwal, John Langford  
HIGHLIGHT: We propose and study a new model for reinforcement learning with rich observations, generalizing contextual bandits to sequential decision making.
- 208, TITLE: Improved Deep Metric Learning with Multi-class N-pair Loss Objective  
<https://papers.nips.cc/paper/6200-improved-deep-metric-learning-with-multi-class-n-pair-loss-objective>  
AUTHORS: Kihyuk Sohn  
HIGHLIGHT: In this paper, we propose to address this problem with a new metric learning objective called multi-class N-pair loss.
- 209, TITLE: Unsupervised Learning of Spoken Language with Visual Context  
<https://papers.nips.cc/paper/6186-unsupervised-learning-of-spoken-language-with-visual-context>  
AUTHORS: David Harwath, Antonio Torralba, James Glass

**HIGHLIGHT:** In this paper, we present a deep neural network model capable of rudimentary spoken language acquisition using untranscribed audio training data, whose only supervision comes in the form of contextually relevant visual images.

210, **TITLE:** Low-Rank Regression with Tensor Responses  
<https://papers.nips.cc/paper/6302-low-rank-regression-with-tensor-responses>

**AUTHORS:** Guillaume Rabusseau, Hachem Kadri

**HIGHLIGHT:** This paper proposes an efficient algorithm (HOLRR) to handle regression tasks where the outputs have a tensor structure.

211, **TITLE:** PAC-Bayesian Theory Meets Bayesian Inference  
<https://papers.nips.cc/paper/6569-pac-bayesian-theory-meets-bayesian-inference>

**AUTHORS:** Pascal Germain, Francis Bach, Alexandre Lacoste, Simon Lacoste-Julien

**HIGHLIGHT:** That is, for the negative log-likelihood loss function, we show that the minimization of PAC-Bayesian generalization bounds maximizes the Bayesian marginal likelihood.

212, **TITLE:** Data Poisoning Attacks on Factorization-Based Collaborative Filtering  
<https://papers.nips.cc/paper/6142-data-poisoning-attacks-on-factorization-based-collaborative-filtering>

**AUTHORS:** Bo Li, Yining Wang, Aarti Singh, Yevgeniy Vorobeychik

**HIGHLIGHT:** We present efficient solutions for two popular factorization-based collaborative filtering algorithms: the alternative minimization formulation and the nuclear norm minimization method.

213, **TITLE:** Learned Region Sparsity and Diversity Also Predicts Visual Attention

<https://papers.nips.cc/paper/6451-learned-region-sparsity-and-diversity-also-predicts-visual-attention>

**AUTHORS:** Zijun Wei, Hossein Adeli, Minh Hoai Nguyen, Greg Zelinsky, Dimitris Samaras

**HIGHLIGHT:** In this paper we incorporate the biologically plausible mechanism of Inhibition of Return into the learned region sparsity model, thereby imposing diversity on the selected regions.

214, **TITLE:** End-to-End Goal-Driven Web Navigation

<https://papers.nips.cc/paper/6064-end-to-end-goal-driven-web-navigation>

**AUTHORS:** Rodrigo Nogueira, Kyunghyun Cho

**HIGHLIGHT:** We propose a goal-driven web navigation as a benchmark task for evaluating an agent with abilities to understand natural language and plan on partially observed environments.

215, **TITLE:** Automated scalable segmentation of neurons from multispectral images

<https://papers.nips.cc/paper/6549-automated-scalable-segmentation-of-neurons-from-multispectral-images>

**AUTHORS:** Uygur S?mb?l, Douglas Roossien, Dawen Cai, Fei Chen, Nicholas Barry, John P. Cunningham, Edward Boyden, Liam Paninski

**HIGHLIGHT:** Here, we propose a method to automate the segmentation of neurons in such (potentially pseudo-colored) images.

216, **TITLE:** Privacy Odometers and Filters: Pay-as-you-Go Composition

<https://papers.nips.cc/paper/6170-privacy-odometers-and-filters-pay-as-you-go-composition>

**AUTHORS:** Ryan M. Rogers, Aaron Roth, Jonathan Ullman, Salil Vadhan

**HIGHLIGHT:** In this paper we initiate the study of adaptive composition in differential privacy when the length of the composition, and the privacy parameters themselves can be chosen adaptively, as a function of the outcome of previously run analyses.

217, **TITLE:** Minimax Estimation of Maximum Mean Discrepancy with Radial Kernels

<https://papers.nips.cc/paper/6483-minimax-estimation-of-maximum-mean-discrepancy-with-radial-kernels>

**AUTHORS:** Ilya O. Tolstikhin, Bharath K. Sriperumbudur, Bernhard Sch?lkopf

**HIGHLIGHT:** In this paper, we present the first known lower bounds for the estimation of MMD based on finite samples.

218, **TITLE:** Adaptive optimal training of animal behavior

<https://papers.nips.cc/paper/6344-adaptive-optimal-training-of-animal-behavior>

**AUTHORS:** Ji Hyun Bak, Jung Yoon Choi, Athena Akrami, Ilana Witten, Jonathan W. Pillow

**HIGHLIGHT:** We develop and test these methods using data collected from rats during training on a two-interval sensory discrimination task.

219, **TITLE:** Hierarchical Object Representation for Open-Ended Object Category Learning and Recognition

<https://papers.nips.cc/paper/6539-hierarchical-object-representation-for-open-ended-object-category-learning-and-recognition>

- AUTHORS: Seyed Hamidreza Kasaei, Ana Maria Tom?, Lu?s Seabra Lopes  
HIGHLIGHT: In particular, we propose an extension of Latent Dirichlet Allocation to learn structural semantic features (i.e. topics) from low-level feature co-occurrences for each category independently.
- 220, TITLE: Relevant sparse codes with variational information bottleneck  
<https://papers.nips.cc/paper/6101-relevant-sparse-codes-with-variational-information-bottleneck>  
AUTHORS: Matthew Chalk, Olivier Marre, Gasper Tkacik  
HIGHLIGHT: Here we propose an approximate variational scheme for maximising a lower bound on the IB objective, analogous to variational EM.
- 221, TITLE: Combinatorial Energy Learning for Image Segmentation  
<https://papers.nips.cc/paper/6595-combinatorial-energy-learning-for-image-segmentation>  
AUTHORS: Jeremy B. Maitin-Shepard, Viren Jain, Michal Januszewski, Peter Li, Pieter Abbeel  
HIGHLIGHT: We propose efficient algorithms for learning deep neural networks to model the energy function, and for local optimization of this energy in the space of supervoxel agglomerations.
- 222, TITLE: Orthogonal Random Features  
<https://papers.nips.cc/paper/6246-orthogonal-random-features>  
AUTHORS: Felix Xinnan X. Yu, Ananda Theertha Suresh, Krzysztof M. Choromanski, Daniel N. Holtmann-Rice, Sanjiv Kumar  
HIGHLIGHT: We present an intriguing discovery related to Random Fourier Features: replacing multiplication by a random Gaussian matrix with multiplication by a properly scaled random orthogonal matrix significantly decreases kernel approximation error.
- 223, TITLE: Fast Active Set Methods for Online Spike Inference from Calcium Imaging  
<https://papers.nips.cc/paper/6505-fast-active-set-methods-for-online-spike-inference-from-calcium-imaging>  
AUTHORS: Johannes Friedrich, Liam Paninski  
HIGHLIGHT: We present a fast online active set method to solve this sparse nonnegative deconvolution problem.
- 224, TITLE: Diffusion-Convolutional Neural Networks  
<https://papers.nips.cc/paper/6212-diffusion-convolutional-neural-networks>  
AUTHORS: James Atwood, Don Towsley  
HIGHLIGHT: We present diffusion-convolutional neural networks (DCNNs), a new model for graph-structured data.
- 225, TITLE: Bayesian latent structure discovery from multi-neuron recordings  
<https://papers.nips.cc/paper/6185-bayesian-latent-structure-discovery-from-multi-neuron-recordings>  
AUTHORS: Scott Linderman, Ryan P. Adams, Jonathan W. Pillow  
HIGHLIGHT: Here we describe new tools for inferring latent structure from simultaneously recorded spike train data using a hierarchical extension of a multi-neuron point process model commonly known as the generalized linear model (GLM).
- 226, TITLE: A Probabilistic Programming Approach To Probabilistic Data Analysis  
<https://papers.nips.cc/paper/6060-a-probabilistic-programming-approach-to-probabilistic-data-analysis>  
AUTHORS: Feras Saad, Vikash K. Mansinghka  
HIGHLIGHT: This paper introduces composable generative population models (CGPMs), a computational abstraction that extends directed graphical models and can be used to describe and compose a broad class of probabilistic data analysis techniques.
- 227, TITLE: A Non-parametric Learning Method for Confidently Estimating Patient's Clinical State and Dynamics  
<https://papers.nips.cc/paper/6454-a-non-parametric-learning-method-for-confidently-estimating-patients-clinical-state-and-dynamics>  
AUTHORS: William Hoiles, Mihaela Van Der Schaar  
HIGHLIGHT: In this paper we construct a non-parametric learning algorithm to estimate the clinical state of a patient.
- 228, TITLE: Inference by Reparameterization in Neural Population Codes  
<https://papers.nips.cc/paper/6476-inference-by-reparameterization-in-neural-population-codes>  
AUTHORS: Rajkumar Vasudeva Raju, Zachary Pitkow  
HIGHLIGHT: Here we present a new general-purpose, biologically-plausible neural implementation of approximate inference.
- 229, TITLE: Tensor Switching Networks  
<https://papers.nips.cc/paper/6516-tensor-switching-networks>  
AUTHORS: Chuan-Yung Tsai, Andrew M. Saxe, Andrew M. Saxe, David Cox

**HIGHLIGHT:** We present a novel neural network algorithm, the Tensor Switching (TS) network, which generalizes the Rectified Linear Unit (ReLU) nonlinearity to tensor-valued hidden units.

230, **TITLE:** Stochastic Gradient Richardson-Romberg Markov Chain Monte Carlo  
<https://papers.nips.cc/paper/6514-stochastic-gradient-richardson-romberg-markov-chain-monte-carlo>  
**AUTHORS:** Alain Durmus, Umut Simsekli, Eric Moulines, Roland Badeau, Gaël RICHARD  
**HIGHLIGHT:** In this study, we propose a novel sampling algorithm that aims to reduce the bias of SG-MCMC while keeping the variance at a reasonable level.

231, **TITLE:** Coordinate-wise Power Method  
<https://papers.nips.cc/paper/6103-coordinate-wise-power-method>  
**AUTHORS:** Qi Lei, Kai Zhong, Inderjit S. Dhillon  
**HIGHLIGHT:** In this paper, we propose a coordinate-wise version of the power method from an optimization viewpoint.

232, **TITLE:** Learning Influence Functions from Incomplete Observations  
<https://papers.nips.cc/paper/6181-learning-influence-functions-from-incomplete-observations>  
**AUTHORS:** Xinran He, Ke Xu, David Kempe, Yan Liu  
**HIGHLIGHT:** We study the problem of learning influence functions under incomplete observations of node activations.

233, **TITLE:** Learning Structured Sparsity in Deep Neural Networks  
<https://papers.nips.cc/paper/6504-learning-structured-sparsity-in-deep-neural-networks>  
**AUTHORS:** Wei Wen, Chunpeng Wu, Yandan Wang, Yiran Chen, Hai Li  
**HIGHLIGHT:** In this work, we propose a Structured Sparsity Learning (SSL) method to regularize the structures (i.e., filters, channels, filter shapes, and layer depth) of DNNs.

234, **TITLE:** Sample Complexity of Automated Mechanism Design  
<https://papers.nips.cc/paper/6351-sample-complexity-of-automated-mechanism-design>  
**AUTHORS:** Maria-Florina F. Balcan, Tuomas Sandholm, Ellen Vitercik  
**HIGHLIGHT:** In this work, we provide the first sample complexity analysis for the standard hierarchy of deterministic combinatorial auction classes used in automated mechanism design.

235, **TITLE:** Short-Dot: Computing Large Linear Transforms Distributedly Using Coded Short Dot Products  
<https://papers.nips.cc/paper/6329-short-dot-computing-large-linear-transforms-distributedly-using-coded-short-dot-products>  
**AUTHORS:** Sanghamitra Dutta, Viveck Cadambe, Pulkit Grover  
**HIGHLIGHT:** In this paper, we propose a novel technique - that we call "Short-Dot" - to introduce redundant computations in a coding theory inspired fashion, for computing linear transforms of long vectors.

236, **TITLE:** Brains on Beats  
<https://papers.nips.cc/paper/6222-brains-on-beats>  
**AUTHORS:** Umut G?l?, Jordy Thielen, Michael Hanke, Marcel van Gerven  
**HIGHLIGHT:** We developed task-optimized deep neural networks (DNNs) that achieved state-of-the-art performance in different evaluation scenarios for automatic music tagging.

237, **TITLE:** Learning Transferrable Representations for Unsupervised Domain Adaptation  
<https://papers.nips.cc/paper/6360-learning-transferrable-representations-for-unsupervised-domain-adaptation>  
**AUTHORS:** Ozan Sener, Hyun Oh Song, Ashutosh Saxena, Silvio Savarese  
**HIGHLIGHT:** In this paper, we propose a unified deep learning framework where the representation, cross domain transformation, and target label inference are all jointly optimized in an end-to-end fashion for unsupervised domain adaptation.

238, **TITLE:** Stochastic Multiple Choice Learning for Training Diverse Deep Ensembles  
<https://papers.nips.cc/paper/6270-stochastic-multiple-choice-learning-for-training-diverse-deep-ensembles>  
**AUTHORS:** Stefan Lee, Senthil Purushwalkam Shiva Prakash, Michael Cogswell, Viresh Ranjan, David Crandall, Dhruv Batra  
**HIGHLIGHT:** In this work, we pose the task of producing multiple outputs as a learning problem over an ensemble of deep networks -- introducing a novel stochastic gradient descent based approach to minimize the loss with respect to an oracle.

239, **TITLE:** Active Learning from Imperfect Labelers  
<https://papers.nips.cc/paper/6162-active-learning-from-imperfect-labelers>  
**AUTHORS:** Songbai Yan, Kamalika Chaudhuri, Tara Javidi

**HIGHLIGHT:** We propose an algorithm which utilizes abstention responses, and analyze its statistical consistency and query complexity under fairly natural assumptions on the noise and abstention rate of the labeler.

240, **TITLE:** Learning to Communicate with Deep Multi-Agent Reinforcement Learning  
<https://papers.nips.cc/paper/6042-learning-to-communicate-with-deep-multi-agent-reinforcement-learning>

**AUTHORS:** Jakob Foerster, Ioannis Alexandros Assael, Nando de Freitas, Shimon Whiteson

**HIGHLIGHT:** We propose two approaches for learning in these domains: Reinforced Inter-Agent Learning (RIAL) and Differentiable Inter-Agent Learning (DIAL).

241, **TITLE:** Value Iteration Networks

<https://papers.nips.cc/paper/6046-value-iteration-networks>

**AUTHORS:** Aviv Tamar, YI WU, Garrett Thomas, Sergey Levine, Pieter Abbeel

**HIGHLIGHT:** We introduce the value iteration network (VIN): a fully differentiable neural network with a 'planning module' embedded within.

242, **TITLE:** Blind Regression: Nonparametric Regression for Latent Variable Models via Collaborative Filtering

<https://papers.nips.cc/paper/6108-blind-regression-nonparametric-regression-for-latent-variable-models-via-collaborative-filtering>

**AUTHORS:** Dogyoon Song, Christina E. Lee, Yihua Li, Devavrat Shah

**HIGHLIGHT:** We introduce the framework of blind regression motivated by matrix completion for recommendation systems: given  $m$  users,  $n$  movies, and a subset of user-movie ratings, the goal is to predict the unobserved user-movie ratings given the data, i.e., to complete the partially observed matrix.

243, **TITLE:** On the Recursive Teaching Dimension of VC Classes

<https://papers.nips.cc/paper/6412-on-the-recursive-teaching-dimension-of-vc-classes>

**AUTHORS:** Xi Chen, Xi Chen, Yu Cheng, Bo Tang

**HIGHLIGHT:** In this paper, we study the quantitative relation between RTD and the well-known learning complexity measure VC dimension (VCD), and improve the best known upper and (worst-case) lower bounds on the recursive teaching dimension with respect to the VC dimension.

244, **TITLE:** InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets

<https://papers.nips.cc/paper/6399-infogan-interpretable-representation-learning-by-information-maximizing-generative-adversarial-nets>

**AUTHORS:** Xi Chen, Yan Duan, Rein Houthoofd, John Schulman, Ilya Sutskever, Pieter Abbeel

**HIGHLIGHT:** This paper describes InfoGAN, an information-theoretic extension to the Generative Adversarial Network that is able to learn disentangled representations in a completely unsupervised manner.

245, **TITLE:** Hardness of Online Sleeping Combinatorial Optimization Problems

<https://papers.nips.cc/paper/6450-hardness-of-online-sleeping-combinatorial-optimization-problems>

**AUTHORS:** Satyen Kale, Chansoo Lee, David Pal

**HIGHLIGHT:** Specifically, we show that the sleeping versions of these problems are at least as hard as PAC learning DNF expressions, a long standing open problem.

246, **TITLE:** Mixed Linear Regression with Multiple Components

<https://papers.nips.cc/paper/6240-mixed-linear-regression-with-multiple-components>

**AUTHORS:** Kai Zhong, Prateek Jain, Inderjit S. Dhillon

**HIGHLIGHT:** In this paper, we study the mixed linear regression (MLR) problem, where the goal is to recover multiple underlying linear models from their unlabeled linear measurements.

247, **TITLE:** Sequential Neural Models with Stochastic Layers

<https://papers.nips.cc/paper/6039-sequential-neural-models-with-stochastic-layers>

**AUTHORS:** Marco Fraccaro, Soren Kaae Snderby, Ulrich Paquet, Ole Winther

**HIGHLIGHT:** This paper introduces stochastic recurrent neural networks which glue a deterministic recurrent neural network and a state space model together to form a stochastic and sequential neural generative model.

248, **TITLE:** Stochastic Gradient Methods for Distributionally Robust Optimization with f-divergences

<https://papers.nips.cc/paper/6040-stochastic-gradient-methods-for-distributionally-robust-optimization-with-f-divergences>

**AUTHORS:** Hongseok Namkoong, John C. Duchi

**HIGHLIGHT:** We develop efficient solution methods for a robust empirical risk minimization problem designed to give calibrated confidence intervals on performance and provide optimal tradeoffs between bias and variance.

- 249, TITLE: Minimizing Quadratic Functions in Constant Time  
<https://papers.nips.cc/paper/6044-minimizing-quadratic-functions-in-constant-time>  
AUTHORS: Kohei Hayashi, Yuichi Yoshida  
HIGHLIGHT: A sampling-based optimization method for quadratic functions is proposed.
- 250, TITLE: Improved Techniques for Training GANs  
<https://papers.nips.cc/paper/6125-improved-techniques-for-training-gans>  
AUTHORS: Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen, Xi Chen  
HIGHLIGHT: We present a variety of new architectural features and training procedures that we apply to the generative adversarial networks (GANs) framework.
- 251, TITLE: DeepMath - Deep Sequence Models for Premise Selection  
<https://papers.nips.cc/paper/6280-deepmath-deep-sequence-models-for-premise-selection>  
AUTHORS: Geoffrey Irving, Christian Szegedy, Alexander A. Alemi, Niklas Een, Francois Chollet, Josef Urban  
HIGHLIGHT: We propose a two stage approach for this task that yields good results for the premise selection task on the Mizar corpus while avoiding the hand-engineered features of existing state-of-the-art models.
- 252, TITLE: Learning Multiagent Communication with Backpropagation  
<https://papers.nips.cc/paper/6398-learning-multiagent-communication-with-backpropagation>  
AUTHORS: Sainbayar Sukhbaatar, arthur szlam, Rob Fergus  
HIGHLIGHT: In this paper we explore a simple neural model, called CommNet, that uses continuous communication for fully cooperative tasks.
- 253, TITLE: Toward Deeper Understanding of Neural Networks: The Power of Initialization and a Dual View on Expressivity  
<https://papers.nips.cc/paper/6427-toward-deeper-understanding-of-neural-networks-the-power-of-initialization-and-a-dual-view-on-expressivity>  
AUTHORS: Amit Daniely, Roy Frostig, Yoram Singer  
HIGHLIGHT: We introduce the notion of a computation skeleton, an acyclic graph that succinctly describes both a family of neural networks and a kernel space.
- 254, TITLE: Learning the Number of Neurons in Deep Networks  
<https://papers.nips.cc/paper/6372-learning-the-number-of-neurons-in-deep-networks>  
AUTHORS: Jose M. Alvarez, Mathieu Salzmann  
HIGHLIGHT: In this paper, we introduce an approach to automatically determining the number of neurons in each layer of a deep network during learning.
- 255, TITLE: Finding significant combinations of features in the presence of categorical covariates  
<https://papers.nips.cc/paper/6345-finding-significant-combinations-of-features-in-the-presence-of-categorical-covariates>  
AUTHORS: Laetitia Papaxanthos, Felipe Llinares-Lopez, Dean Bodenham, Karsten Borgwardt  
HIGHLIGHT: We propose the Fast Automatic Conditional Search (FACS) algorithm, a significant discriminative itemset mining method which conditions on categorical covariates and only scales as  $O(k \log k)$ , where  $k$  is the number of states of the categorical covariate.
- 256, TITLE: Examples are not enough, learn to criticize! Criticism for Interpretability  
<https://papers.nips.cc/paper/6300-examples-are-not-enough-learn-to-criticize-criticism-for-interpretability>  
AUTHORS: Been Kim, Rajiv Khanna, Oluwasanmi O. Koyejo  
HIGHLIGHT: Example-based explanations are widely used in the effort to improve the interpretability of highly complex distributions.
- 257, TITLE: Optimistic Bandit Convex Optimization  
<https://papers.nips.cc/paper/6429-optimistic-bandit-convex-optimization>  
AUTHORS: Scott Yang, Mehryar Mohri  
HIGHLIGHT: We introduce the general and powerful scheme of predicting information re-use in optimization algorithms.
- 258, TITLE: Safe Policy Improvement by Minimizing Robust Baseline Regret  
<https://papers.nips.cc/paper/6294-safe-policy-improvement-by-minimizing-robust-baseline-regret>  
AUTHORS: Mohammad Ghavamzadeh, Marek Petrik, Yinlam Chow  
HIGHLIGHT: In this paper, we develop and analyze a new model-based approach to compute a safe policy when we have access to an inaccurate dynamics model of the system with known accuracy guarantees.

- 259, TITLE: Graphons, mergeons, and so on!  
<https://papers.nips.cc/paper/6089-graphons-mergeons-and-so-on>  
AUTHORS: Justin Eldridge, Mikhail Belkin, Yusu Wang  
HIGHLIGHT: In this work we develop a theory of hierarchical clustering for graphs.
- 260, TITLE: Hierarchical Clustering via Spreading Metrics  
<https://papers.nips.cc/paper/6325-hierarchical-clustering-via-spreading-metrics>  
AUTHORS: Aurko Roy, Sebastian Pokutta  
HIGHLIGHT: We study the cost function for hierarchical clusterings introduced by [Dasgupta, 2015] where hierarchies are treated as first-class objects rather than deriving their cost from projections into flat clusters.
- 261, TITLE: Learning Bayesian networks with ancestral constraints  
<https://papers.nips.cc/paper/6386-learning-bayesian-networks-with-ancestral-constraints>  
AUTHORS: Eunice Yuh-Jie Chen, Yujia Shen, Arthur Choi, Adnan Darwiche  
HIGHLIGHT: We consider the problem of learning Bayesian networks optimally, when subject to background knowledge in the form of ancestral constraints.
- 262, TITLE: Pruning Random Forests for Prediction on a Budget  
<https://papers.nips.cc/paper/6250-pruning-random-forests-for-prediction-on-a-budget>  
AUTHORS: Feng Nan, Joseph Wang, Venkatesh Saligrama  
HIGHLIGHT: We propose to prune a random forest (RF) for resource-constrained prediction.
- 263, TITLE: Clustering with Bregman Divergences: an Asymptotic Analysis  
<https://papers.nips.cc/paper/6550-clustering-with-bregman-divergences-an-asymptotic-analysis>  
AUTHORS: Chaoyue Liu, Mikhail Belkin  
HIGHLIGHT: In this paper we analyze theoretical properties of Bregman clustering when the number of the clusters  $k$  is large.
- 264, TITLE: Variational Autoencoder for Deep Learning of Images, Labels and Captions  
<https://papers.nips.cc/paper/6528-variational-autoencoder-for-deep-learning-of-images-labels-and-captions>  
AUTHORS: Yunchen Pu, Zhe Gan, Ricardo Henao, Xin Yuan, Chunyuan Li, Andrew Stevens, Lawrence Carin  
HIGHLIGHT: Variational Autoencoder for Deep Learning of Images, Labels and Captions
- 265, TITLE: Review Networks for Caption Generation  
<https://papers.nips.cc/paper/6167-review-networks-for-caption-generation>  
AUTHORS: Zhilin Yang, Ye Yuan, Yuexin Wu, William W. Cohen, Ruslan R. Salakhutdinov  
HIGHLIGHT: We propose a novel extension of the encoder-decoder framework, called a review network.
- 266, TITLE: Stein Variational Gradient Descent: A General Purpose Bayesian Inference Algorithm  
<https://papers.nips.cc/paper/6338-stein-variational-gradient-descent-a-general-purpose-bayesian-inference-algorithm>  
AUTHORS: Qiang Liu, Dilin Wang  
HIGHLIGHT: We propose a general purpose variational inference algorithm that forms a natural counterpart of gradient descent for optimization.
- 267, TITLE: A Bio-inspired Redundant Sensing Architecture  
<https://papers.nips.cc/paper/6564-a-bio-inspired-redundant-sensing-architecture>  
AUTHORS: Anh Tuan Nguyen, Jian Xu, Zhi Yang  
HIGHLIGHT: In this work, we argue that redundancy is one of the critical characteristics for such superior performance.
- 268, TITLE: Contextual semibandits via supervised learning oracles  
<https://papers.nips.cc/paper/6513-contextual-semibandits-via-supervised-learning-oracles>  
AUTHORS: Akshay Krishnamurthy, Alekh Agarwal, Miro Dudik  
HIGHLIGHT: We study an online decision making problem where on each round a learner chooses a list of items based on some side information, receives a scalar feedback value for each individual item, and a reward that is linearly related to this feedback.
- 269, TITLE: Blind Attacks on Machine Learners  
<https://papers.nips.cc/paper/6482-blind-attacks-on-machine-learners>  
AUTHORS: Alex Beatson, Zhaoran Wang, Han Liu

**HIGHLIGHT:** We study the potential of a “blind attacker” to provably limit a learner’s performance by data injection attack without observing the learner’s training set or any parameter of the distribution from which it is drawn.

270, **TITLE:** Universal Correspondence Network  
<https://papers.nips.cc/paper/6487-universal-correspondence-network>  
**AUTHORS:** Christopher B. Choy, JunYoung Gwak, Silvio Savarese, Manmohan Chandraker  
**HIGHLIGHT:** We present a deep learning framework for accurate visual correspondences and demonstrate its effectiveness for both geometric and semantic matching, spanning across rigid motions to intra-class shape or appearance variations.

271, **TITLE:** Satisfying Real-world Goals with Dataset Constraints  
<https://papers.nips.cc/paper/6316-satisfying-real-world-goals-with-dataset-constraints>  
**AUTHORS:** Gabriel Goh, Andrew Cotter, Maya Gupta, Michael P. Friedlander  
**HIGHLIGHT:** In this paper we propose handling multiple goals on multiple datasets by training with dataset constraints, using the ramp penalty to accurately quantify costs, and present an efficient algorithm to approximately optimize the resulting non-convex constrained optimization problem.

272, **TITLE:** Deep Learning for Predicting Human Strategic Behavior  
<https://papers.nips.cc/paper/6509-deep-learning-for-predicting-human-strategic-behavior>  
**AUTHORS:** Jason S. Hartford, James R. Wright, Kevin Leyton-Brown  
**HIGHLIGHT:** In this work, we present an alternative, a deep learning approach that automatically performs cognitive modeling without relying on such expert knowledge.

273, **TITLE:** Phased Exploration with Greedy Exploitation in Stochastic Combinatorial Partial Monitoring Games  
<https://papers.nips.cc/paper/6198-phased-exploration-with-greedy-exploitation-in-stochastic-combinatorial-partial-monitoring-games>  
**AUTHORS:** Sougata Chaudhuri, Ambuj Tewari  
**HIGHLIGHT:** Recently, a general model of combinatorial partial monitoring (CPM) games was proposed [Lincombinatorial2014], where the learner’s action space can be exponentially large and adversary samples its moves from a bounded, continuous space, according to a fixed distribution.

274, **TITLE:** Eliciting Categorical Data for Optimal Aggregation  
<https://papers.nips.cc/paper/6237-eliciting-categorical-data-for-optimal-aggregation>  
**AUTHORS:** Chien-Ju Ho, Rafael Frongillo, Yiling Chen  
**HIGHLIGHT:** In this paper, we develop a Bayesian model, wherein agents have differing quality of information, but also respond to incentives.

275, **TITLE:** Measuring the reliability of MCMC inference with bidirectional Monte Carlo  
<https://papers.nips.cc/paper/6290-measuring-the-reliability-of-mcmc-inference-with-bidirectional-monte-carlo>  
**AUTHORS:** Roger B. Grosse, Siddharth Ancha, Daniel M. Roy  
**HIGHLIGHT:** In this work, we extend the recently introduced bidirectional Monte Carlo technique to evaluate MCMC-based posterior inference algorithms.

276, **TITLE:** Breaking the Bandwidth Barrier: Geometrical Adaptive Entropy Estimation  
<https://papers.nips.cc/paper/6299-breaking-the-bandwidth-barrier-geometrical-adaptive-entropy-estimation>  
**AUTHORS:** Weihao Gao, Sewoong Oh, Pramod Viswanath  
**HIGHLIGHT:** In this paper we combine both these approaches to design new estimators of entropy and mutual information that strongly outperform all state of the art methods.

277, **TITLE:** Selective inference for group-sparse linear models  
<https://papers.nips.cc/paper/6437-selective-inference-for-group-sparse-linear-models>  
**AUTHORS:** Fan Yang, Rina Foygel Barber, Prateek Jain, John Lafferty  
**HIGHLIGHT:** We develop tools for selective inference in the setting of group sparsity, including the construction of confidence intervals and p-values for testing selected groups of variables.

278, **TITLE:** Graph Clustering: Block-models and model free results  
<https://papers.nips.cc/paper/6140-graph-clustering-block-models-and-model-free-results>  
**AUTHORS:** Yali Wan, Marina Meila  
**HIGHLIGHT:** In this paper, we propose a framework, in which we obtain “correctness” guarantees without assuming the data comes from a model.

279, **TITLE:** Maximizing Influence in an Ising Network: A Mean-Field Optimal Solution

- <https://papers.nips.cc/paper/6119-maximizing-influence-in-an-ising-network-a-mean-field-optimal-solution>  
AUTHORS: Christopher Lynn, Daniel D. Lee  
HIGHLIGHT: In this paper, we consider an alternate model that treats individual opinions as spins in an Ising system at dynamic equilibrium.
- 280, TITLE: Hypothesis Testing in Unsupervised Domain Adaptation with Applications in Alzheimer's Disease  
<https://papers.nips.cc/paper/6209-hypothesis-testing-in-unsupervised-domain-adaptation-with-applications-in-alzheimers-disease>  
AUTHORS: Hao Zhou, Vamsi K. Ithapu, Sathya Narayanan Ravi, Vikas Singh, Grace Wahba, Sterling C. Johnson  
HIGHLIGHT: Our goal is to perform a statistical test checking if  $P_{\text{source}} = P_{\text{target}}$  while removing the distortions induced by the transformations.
- 281, TITLE: Geometric Dirichlet Means Algorithm for topic inference  
<https://papers.nips.cc/paper/6332-geometric-dirichlet-means-algorithm-for-topic-inference>  
AUTHORS: Mikhail Yurochkin, XuanLong Nguyen  
HIGHLIGHT: We propose a geometric algorithm for topic learning and inference that is built on the convex geometry of topics arising from the Latent Dirichlet Allocation (LDA) model and its nonparametric extensions.
- 282, TITLE: Structured Prediction Theory Based on Factor Graph Complexity  
<https://papers.nips.cc/paper/6485-structured-prediction-theory-based-on-factor-graph-complexity>  
AUTHORS: Corinna Cortes, Vitaly Kuznetsov, Mehryar Mohri, Scott Yang  
HIGHLIGHT: We present new learning bounds for this advanced setting, which we use to devise two new algorithms, `Voted Conditional Random Field` (VCRF) and `Voted Structured Boosting` (StructBoost).
- 283, TITLE: Improved Dropout for Shallow and Deep Learning  
<https://papers.nips.cc/paper/6561-improved-dropout-for-shallow-and-deep-learning>  
AUTHORS: Zhe Li, Boqing Gong, Tianbao Yang  
HIGHLIGHT: In this paper, we propose to use multinomial sampling for dropout, i.e., sampling features or neurons according to a multinomial distribution with different probabilities for different features/neurons.
- 284, TITLE: Constraints Based Convex Belief Propagation  
<https://papers.nips.cc/paper/6219-constraints-based-convex-belief-propagation>  
AUTHORS: Yaniv Tenzer, Alex Schwing, Kevin Gimpel, Tamir Hazan  
HIGHLIGHT: In this paper we suggest to tackle consistency by incorporating constraints on beliefs.
- 285, TITLE: Error Analysis of Generalized Nyström Kernel Regression  
<https://papers.nips.cc/paper/6602-error-analysis-of-generalized-nyström-kernel-regression>  
AUTHORS: Hong Chen, Haifeng Xia, Heng Huang, Weidong Cai  
HIGHLIGHT: In this paper, we consider the generalized Nyström kernel regression (GNKR) with  $\ell_2$  coefficient regularization, where the kernel just requires the continuity and boundedness.
- 286, TITLE: A Probabilistic Framework for Deep Learning  
<https://papers.nips.cc/paper/6231-a-probabilistic-framework-for-deep-learning>  
AUTHORS: Ankit B. Patel, Minh Tan Nguyen, Richard Baraniuk  
HIGHLIGHT: We develop a probabilistic framework for deep learning based on the Deep Rendering Mixture Model (DRMM), a new generative probabilistic model that explicitly capture variations in data due to latent task nuisance variables.
- 287, TITLE: General Tensor Spectral Co-clustering for Higher-Order Data  
<https://papers.nips.cc/paper/6376-general-tensor-spectral-co-clustering-for-higher-order-data>  
AUTHORS: Tao Wu, Austin R. Benson, David F. Gleich  
HIGHLIGHT: Spectral clustering and co-clustering are well-known techniques in data analysis, and recent work has extended spectral clustering to square, symmetric tensors and hypermatrices derived from a network.
- 288, TITLE: Cyclades: Conflict-free Asynchronous Machine Learning  
<https://papers.nips.cc/paper/6604-cyclades-conflict-free-asynchronous-machine-learning>  
AUTHORS: Xinghao Pan, Maximilian Lam, Stephen Tu, Dimitris Papailiopoulos, Ce Zhang, Michael I. Jordan, Kannan Ramchandran, Christopher R?  
HIGHLIGHT: We present Cyclades, a general framework for parallelizing stochastic optimization algorithms in a shared memory setting.
- 289, TITLE: Single Pass PCA of Matrix Products

- <https://papers.nips.cc/paper/6075-single-pass-pca-of-matrix-products>  
AUTHORS: Shanshan Wu, Srinadh Bhojanapalli, Sujay Sanghavi, Alexandros G. Dimakis  
HIGHLIGHT: In this paper we present a new algorithm for computing a low rank approximation of the product  $A^T B S$  by taking only a single pass of the two matrices  $A S$  and  $B S$ .
- 290, TITLE: Stochastic Variational Deep Kernel Learning  
<https://papers.nips.cc/paper/6426-stochastic-variational-deep-kernel-learning>  
AUTHORS: Andrew G. Wilson, Zhiting Hu, Ruslan R. Salakhutdinov, Eric P. Xing  
HIGHLIGHT: We propose a novel deep kernel learning model and stochastic variational inference procedure which generalizes deep kernel learning approaches to enable classification, multi-task learning, additive covariance structures, and stochastic gradient training.
- 291, TITLE: Interaction Screening: Efficient and Sample-Optimal Learning of Ising Models  
<https://papers.nips.cc/paper/6375-interaction-screening-efficient-and-sample-optimal-learning-of-ising-models>  
AUTHORS: Marc Vuffray, Sidhant Misra, Andrey Lokhov, Michael Chertkov  
HIGHLIGHT: We consider the problem of learning the underlying graph of an unknown Ising model on  $p$  spins from a collection of i.i.d. samples generated from the model.
- 292, TITLE: Long-term Causal Effects via Behavioral Game Theory  
<https://papers.nips.cc/paper/6059-long-term-causal-effects-via-behavioral-game-theory>  
AUTHORS: Panagiotis Toulis, David C. Parkes  
HIGHLIGHT: % Here, we formalize a framework to define and estimate long-term causal effects of policy changes in multiagent economies.
- 293, TITLE: Measuring Neural Net Robustness with Constraints  
<https://papers.nips.cc/paper/6339-measuring-neural-net-robustness-with-constraints>  
AUTHORS: Osbert Bastani, Yani Ioannou, Leonidas Lampropoulos, Dimitrios Vytiniotis, Aditya Nori, Antonio Criminisi  
HIGHLIGHT: We propose metrics for measuring the robustness of a neural net and devise a novel algorithm for approximating these metrics based on an encoding of robustness as a linear program.
- 294, TITLE: Reshaped Wirtinger Flow for Solving Quadratic System of Equations  
<https://papers.nips.cc/paper/6319-reshaped-wirtinger-flow-for-solving-quadratic-system-of-equations>  
AUTHORS: Huishuai Zhang, Yingbin Liang  
HIGHLIGHT: We study the problem of recovering a vector  $\mathbf{x}$  in  $\mathbb{R}^n$  from its magnitude measurements  $y_i = \|\mathbf{a}_i \cdot \mathbf{x}\|$ ,  $i=1, \dots, m$ .
- 295, TITLE: Nearly Isometric Embedding by Relaxation  
<https://papers.nips.cc/paper/6535-nearly-isometric-embedding-by-relaxation>  
AUTHORS: James McQueen, Marina Meila, Dominique Joncas  
HIGHLIGHT: This paper proposes an embedding algorithm that overcomes this problem.
- 296, TITLE: Probabilistic Inference with Generating Functions for Poisson Latent Variable Models  
<https://papers.nips.cc/paper/6587-probabilistic-inference-with-generating-functions-for-poisson-latent-variable-models>  
AUTHORS: Kevin Winner, Daniel R. Sheldon  
HIGHLIGHT: Standard exact inference techniques such as variable elimination and belief propagation do not apply to these models because the latent variables have countably infinite support.
- 297, TITLE: Causal meets Submodular: Subset Selection with Directed Information  
<https://papers.nips.cc/paper/6384-causal-meets-submodular-subset-selection-with-directed-information>  
AUTHORS: Yuxun Zhou, Costas J. Spanos  
HIGHLIGHT: To substantiate the idea of approximate submodularity, we introduce a novel quantity, namely submodularity index (SmI), for general set functions.
- 298, TITLE: Depth from a Single Image by Harmonizing Overcomplete Local Network Predictions  
<https://papers.nips.cc/paper/6510-depth-from-a-single-image-by-harmonizing-overcomplete-local-network-predictions>  
AUTHORS: Ayan Chakrabarti, Jingyu Shao, Greg Shakhnarovich  
HIGHLIGHT: We approach the problem of monocular depth estimation by using a neural network to produce a mid-level representation that summarizes these cues.
- 299, TITLE: Deep Neural Networks with Inexact Matching for Person Re-Identification

<https://papers.nips.cc/paper/6367-deep-neural-networks-with-inexact-matching-for-person-re-identification>

AUTHORS: Arulkumar Subramaniam, Moitreyia Chatterjee, Anurag Mittal  
HIGHLIGHT: In this work, we propose two CNN-based architectures for Person Re-Identification.

300, TITLE: Global Analysis of Expectation Maximization for Mixtures of Two Gaussians  
<https://papers.nips.cc/paper/6047-global-analysis-of-expectation-maximization-for-mixtures-of-two-gaussians>

AUTHORS: Ji Xu, Daniel J. Hsu, Arian Maleki  
HIGHLIGHT: This article addresses this disconnect between the statistical principles behind EM and its algorithmic properties.

301, TITLE: Estimating the class prior and posterior from noisy positives and unlabeled data  
<https://papers.nips.cc/paper/6168-estimating-the-class-prior-and-posterior-from-noisy-positives-and-unlabeled-data>

AUTHORS: Shantanu Jain, Martha White, Predrag Radivojac  
HIGHLIGHT: We develop a classification algorithm for estimating posterior distributions from positive-unlabeled data, that is robust to noise in the positive labels and effective for high-dimensional data.

302, TITLE: Kronecker Determinantal Point Processes  
<https://papers.nips.cc/paper/6296-kronecker-determinantal-point-processes>

AUTHORS: Zelda E. Mariet, Suvrit Sra  
HIGHLIGHT: We enable efficient sampling and learning for DPPs by introducing KronDPP, a DPP model whose kernel matrix decomposes as a tensor product of multiple smaller kernel matrices.

303, TITLE: Finite Sample Prediction and Recovery Bounds for Ordinal Embedding  
<https://papers.nips.cc/paper/6554-finite-sample-prediction-and-recovery-bounds-for-ordinal-embedding>

AUTHORS: Lalit Jain, Kevin G. Jamieson, Rob Nowak  
HIGHLIGHT: To account for errors and variation in judgments, we consider the noisy situation in which the given constraints are independently corrupted by reversing the correct constraint with some probability.

304, TITLE: Feature-distributed sparse regression: a screen-and-clean approach  
<https://papers.nips.cc/paper/6187-feature-distributed-sparse-regression-a-screen-and-clean-approach>

AUTHORS: Jiyang Yang, Michael W. Mahoney, Michael Saunders, Yuekai Sun  
HIGHLIGHT: We propose an algorithm to distributed sparse regression when the data is partitioned by features rather than samples.

305, TITLE: Learning Bound for Parameter Transfer Learning  
<https://papers.nips.cc/paper/6136-learning-bound-for-parameter-transfer-learning>

AUTHORS: Wataru Kumagai  
HIGHLIGHT: We consider a transfer-learning problem by using the parameter transfer approach, where a suitable parameter of feature mapping is learned through one task and applied to another objective task.

306, TITLE: Learning under uncertainty: a comparison between R-W and Bayesian approach  
<https://papers.nips.cc/paper/6409-learning-under-uncertainty-a-comparison-between-r-w-and-bayesian-approach>

AUTHORS: He Huang, Martin Paulus  
HIGHLIGHT: To examine the underlying computational principles that guide different learning behavior in an uncertain environment, we compared an R-W model and a Bayesian approach in a visual search task with different volatility levels.

307, TITLE: Bi-Objective Online Matching and Submodular Allocations  
<https://papers.nips.cc/paper/6085-bi-objective-online-matching-and-submodular-allocations>

AUTHORS: Hossein Esfandiari, Nitish Korula, Vahab Mirrokni  
HIGHLIGHT: In this paper, we give the first results for bi-objective online submodular optimization, providing almost matching upper and lower bounds for allocating items to agents with two submodular value functions.

308, TITLE: Quantized Random Projections and Non-Linear Estimation of Cosine Similarity  
<https://papers.nips.cc/paper/6492-quantized-random-projections-and-non-linear-estimation-of-cosine-similarity>

AUTHORS: Ping Li, Michael Mitzenmacher, Martin Slawski  
HIGHLIGHT: In the present paper, we consider the problem of estimating cosine similarities when the projected data undergo scalar quantization to  $b$  bits.

309, TITLE: The non-convex Burer-Monteiro approach works on smooth semidefinite programs  
<https://papers.nips.cc/paper/6517-the-non-convex-burer-monteiro-approach-works-on-smooth-semidefinite-programs>

AUTHORS: Nicolas Boumal, Vlad Voroninski, Afonso Bandeira  
HIGHLIGHT: In this paper, we consider a class of SDP's which includes applications such as max-cut, community detection in the stochastic block model, robust PCA, phase retrieval and synchronization of rotations.

310, TITLE: Dimensionality Reduction of Massive Sparse Datasets Using Coresets  
<https://papers.nips.cc/paper/6596-dimensionality-reduction-of-massive-sparse-datasets-using-coresets>

AUTHORS: Dan Feldman, Mikhail Volkov, Daniela Rus  
HIGHLIGHT: In this paper we present a practical solution with performance guarantees to the problem of dimensionality reduction for very large scale sparse matrices.

311, TITLE: Using Social Dynamics to Make Individual Predictions: Variational Inference with a Stochastic Kinetic Model  
<https://papers.nips.cc/paper/6453-using-social-dynamics-to-make-individual-predictions-variational-inference-with-a-stochastic-kinetic-model>

AUTHORS: Zhen Xu, Wen Dong, Sargur N. Srihari  
HIGHLIGHT: In order to cope with dynamic interactions among a large number of individuals, we introduce the stochastic kinetic model to capture adaptive transition probabilities and propose an efficient variational inference algorithm the complexity of which grows linearly — rather than exponentially— with the number of individuals.

312, TITLE: Supervised learning through the lens of compression  
<https://papers.nips.cc/paper/6490-supervised-learning-through-the-lens-of-compression>

AUTHORS: Ofir David, Shay Moran, Amir Yehudayoff  
HIGHLIGHT: We use the compressibility-learnability equivalence to show that (i) for multiclass categorization, PAC and agnostic PAC learnability are equivalent, and (ii) to derive a compactness theorem for learnability.

313, TITLE: Generative Shape Models: Joint Text Recognition and Segmentation with Very Little Training Data  
<https://papers.nips.cc/paper/6071-generative-shape-models-joint-text-recognition-and-segmentation-with-very-little-training-data>

AUTHORS: Xinghua Lou, Ken Kansky, Wolfgang Leirach, CC Laan, Bhaskara Marthi, D. Phoenix, Dileep George  
HIGHLIGHT: We demonstrate that a generative model for object shapes can achieve state of the art results on challenging scene text recognition tasks, and with orders of magnitude fewer training images than required for competing discriminative methods.

314, TITLE: Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections  
<https://papers.nips.cc/paper/6172-image-restoration-using-very-deep-convolutional-encoder-decoder-networks-with-symmetric-skip-connections>

AUTHORS: Xiaojiao Mao, Chunhua Shen, Yu-Bin Yang  
HIGHLIGHT: In this paper, we propose a very deep fully convolutional encoding-decoding framework for image restoration such as denoising and super-resolution.

315, TITLE: Object based Scene Representations using Fisher Scores of Local Subspace Projections  
<https://papers.nips.cc/paper/6343-object-based-scene-representations-using-fisher-scores-of-local-subspace-projections>

AUTHORS: Mandar D. Dixit, Nuno Vasconcelos  
HIGHLIGHT: Several works have shown that deep CNN classifiers can be easily transferred across datasets, e.g. the transfer of a CNN trained to recognize objects on ImageNET to an object detector on Pascal VOC.

316, TITLE: Active Learning with Oracle Epiphany  
<https://papers.nips.cc/paper/6155-active-learning-with-oracle-epiphany>

AUTHORS: Tzu-Kuo Huang, Lihong Li, Ara Vartanian, Saleema Amershi, Jerry Zhu  
HIGHLIGHT: We present a theoretical analysis of active learning with more realistic interactions with human oracles.

317, TITLE: Statistical Inference for Pairwise Graphical Models Using Score Matching  
<https://papers.nips.cc/paper/6530-statistical-inference-for-pairwise-graphical-models-using-score-matching>

AUTHORS: Ming Yu, Mladen Kolar, Varun Gupta  
HIGHLIGHT: In this paper, we propose a novel estimator for edge parameters for pairwise graphical models based on Hyvärinen scoring rule.

318, TITLE: Improved Error Bounds for Tree Representations of Metric Spaces  
<https://papers.nips.cc/paper/6431-improved-error-bounds-for-tree-representations-of-metric-spaces>

AUTHORS: Samir Chowdhury, Facundo M. Moli, Zane T. Smith  
HIGHLIGHT: Existing algorithms for embedding metric spaces into tree metrics provide distortion bounds depending on cardinality.

- 319, TITLE: Can Peripheral Representations Improve Clutter Metrics on Complex Scenes?  
<https://papers.nips.cc/paper/6404-can-peripheral-representations-improve-clutter-metrics-on-complex-scenes>  
AUTHORS: Arturo Deza, Miguel Eckstein  
HIGHLIGHT: Here, we introduce a new foveated clutter model to predict the detrimental effects in target search utilizing a forced fixation search task.
- 320, TITLE: On Multiplicative Integration with Recurrent Neural Networks  
<https://papers.nips.cc/paper/6215-on-multiplicative-integration-with-recurrent-neural-networks>  
AUTHORS: Yuhuai Wu, Saizheng Zhang, Ying Zhang, Yoshua Bengio, Ruslan R. Salakhutdinov  
HIGHLIGHT: We introduce a general simple structural design called "Multiplicative Integration" (MI) to improve recurrent neural networks (RNNs).
- 321, TITLE: Learning HMMs with Nonparametric Emissions via Spectral Decompositions of Continuous Matrices  
<https://papers.nips.cc/paper/6086-learning-hmms-with-nonparametric-emissions-via-spectral-decompositions-of-continuous-matrices>  
AUTHORS: Kirthevasan Kandasamy, Maruan Al-Shedivat, Eric P. Xing  
HIGHLIGHT: In this paper, we study the estimation of an  $M$ -state hidden Markov model (HMM) with only smoothness assumptions, such as Hölderian conditions, on the emission densities.
- 322, TITLE: Regret Bounds for Non-decomposable Metrics with Missing Labels  
<https://papers.nips.cc/paper/6178-regret-bounds-for-non-decomposable-metrics-with-missing-labels>  
AUTHORS: Nagarajan Natarajan, Prateek Jain  
HIGHLIGHT: To this end, we propose a generic framework that given a performance metric  $\Psi$ , can devise a regularized objective function and a threshold such that all the values in the predicted score vector above and only above the threshold are selected to be positive.
- 323, TITLE: Robust k-means: a Theoretical Revisit  
<https://papers.nips.cc/paper/6126-robust-k-means-a-theoretical-revisit>  
AUTHORS: ALEXANDROS GEORGOGIANNIS  
HIGHLIGHT: In this work, we present a theoretical analysis of the robustness and consistency properties of a variant of the classical quadratic k-means algorithm, the robust k-means, which borrows ideas from outlier detection in regression.
- 324, TITLE: Bayesian optimization for automated model selection  
<https://papers.nips.cc/paper/6466-bayesian-optimization-for-automated-model-selection>  
AUTHORS: Gustavo Malkomes, Charles Schaff, Roman Garnett  
HIGHLIGHT: We present a sophisticated method for automatically searching for an appropriate kernel from an infinite space of potential choices.
- 325, TITLE: A Probabilistic Model of Social Decision Making based on Reward Maximization  
<https://papers.nips.cc/paper/6537-a-probabilistic-model-of-social-decision-making-based-on-reward-maximization>  
AUTHORS: Koosha Khalvati, Seongmin A. Park, Jean-Claude Dreher, Rajesh P. Rao  
HIGHLIGHT: We employ the framework of partially observable Markov decision processes (POMDPs) to model human decision making in a social context, focusing specifically on the volunteer's dilemma in a version of the classic Public Goods Game.
- 326, TITLE: Balancing Suspense and Surprise: Timely Decision Making with Endogenous Information Acquisition  
<https://papers.nips.cc/paper/6062-balancing-suspense-and-surprise-timely-decision-making-with-endogenous-information-acquisition>  
AUTHORS: Ahmed M. Alaa, Mihaela van der Schaar  
HIGHLIGHT: We develop a Bayesian model for decision-making under time pressure with endogenous information acquisition.
- 327, TITLE: Fast and Flexible Monotonic Functions with Ensembles of Lattices  
<https://papers.nips.cc/paper/6377-fast-and-flexible-monotonic-functions-with-ensembles-of-lattices>  
AUTHORS: Mahdi Milani Fard, Kevin Canini, Andrew Cotter, Jan Pfeifer, Maya Gupta  
HIGHLIGHT: For many machine learning problems, there are some inputs that are known to be positively (or negatively) related to the output, and in such cases training the model to respect that monotonic relationship can provide regularization, and makes the model more interpretable.
- 328, TITLE: Conditional Generative Moment-Matching Networks  
<https://papers.nips.cc/paper/6255-conditional-generative-moment-matching-networks>  
AUTHORS: Yong Ren, Jun Zhu, Jialian Li, Yucen Luo

**HIGHLIGHT:** In this paper, we present conditional generative moment-matching networks (CGMMN), which learn a conditional distribution given some input variables based on a conditional maximum mean discrepancy (CMMD) criterion.

329, **TITLE:** Stochastic Gradient MCMC with Stale Gradients

<https://papers.nips.cc/paper/6359-stochastic-gradient-mcmc-with-stale-gradients>

**AUTHORS:** Changyou Chen, Nan Ding, Chunyuan Li, Yizhe Zhang, Lawrence Carin

**HIGHLIGHT:** In this paper we develop theory to show that while the bias and MSE of an SG-MCMC algorithm depend on the staleness of stochastic gradients, its estimation variance (relative to the expected estimate, based on a prescribed number of samples) is independent of it.

330, **TITLE:** Composing graphical models with neural networks for structured representations and fast inference

<https://papers.nips.cc/paper/6379-composing-graphical-models-with-neural-networks-for-structured-representations-and-fast-inference>

**AUTHORS:** Matthew J. Johnson, David K. Duvenaud, Alex Wiltschko, Ryan P. Adams, Sandeep R. Datta

**HIGHLIGHT:** We propose a general modeling and inference framework that combines the complementary strengths of probabilistic graphical models and deep learning methods.

331, **TITLE:** Noise-Tolerant Life-Long Matrix Completion via Adaptive Sampling

<https://papers.nips.cc/paper/6580-noise-tolerant-life-long-matrix-completion-via-adaptive-sampling>

**AUTHORS:** Maria-Florina F. Balcan, Hongyang Zhang

**HIGHLIGHT:** In this work, we give algorithms achieving strong guarantee under two realistic noise models.

332, **TITLE:** Combinatorial semi-bandit with known covariance

<https://papers.nips.cc/paper/6137-combinatorial-semi-bandit-with-known-covariance>

**AUTHORS:** Romy Degenne, Vianney Perchet

**HIGHLIGHT:** We introduce a way to quantify the dependency structure of the problem and design an algorithm that adapts to it.

333, **TITLE:** Matrix Completion has No Spurious Local Minimum

<https://papers.nips.cc/paper/6048-matrix-completion-has-no-spurious-local-minimum>

**AUTHORS:** Rong Ge, Jason D. Lee, Tengyu Ma

**HIGHLIGHT:** Therefore, many popular optimization algorithms such as (stochastic) gradient descent can provably solve matrix completion with `arbitrary` initialization in polynomial time.

334, **TITLE:** The Multiscale Laplacian Graph Kernel

<https://papers.nips.cc/paper/6135-the-multiscale-laplacian-graph-kernel>

**AUTHORS:** Risi Kondor, Horace Pan

**HIGHLIGHT:** In contrast, by building a hierarchy of nested subgraphs, the Multiscale Laplacian Graph kernels (MLG kernels) that we define in this paper can account for structure at a range of different scales.

335, **TITLE:** Adaptive Averaging in Accelerated Descent Dynamics

<https://papers.nips.cc/paper/6553-adaptive-averaging-in-accelerated-descent-dynamics>

**AUTHORS:** Walid Krichene, Alexandre Bayen, Peter L. Bartlett

**HIGHLIGHT:** We study accelerated descent dynamics for constrained convex optimization.

336, **TITLE:** Sub-sampled Newton Methods with Non-uniform Sampling

<https://papers.nips.cc/paper/6037-sub-sampled-newton-methods-with-non-uniform-sampling>

**AUTHORS:** Peng Xu, Jiyan Yang, Farbod Roosta-Khorasani, Christopher R?, Michael W. Mahoney

**HIGHLIGHT:** We propose randomized Newton-type algorithms that exploit `non-uniform` sub-sampling of  $\{\nabla^2 f_i(w)\}_{i=1}^n$ , as well as inexact updates, as means to reduce the computational complexity, and are applicable to a wide range of problems in machine learning.

337, **TITLE:** Stochastic Gradient Geodesic MCMC Methods

<https://papers.nips.cc/paper/6282-stochastic-gradient-geodesic-mcmc-methods>

**AUTHORS:** Chang Liu, Jun Zhu, Yang Song

**HIGHLIGHT:** We propose two stochastic gradient MCMC methods for sampling from Bayesian posterior distributions defined on Riemann manifolds with a known geodesic flow, e.g. hyperspheres.

338, **TITLE:** Variational Bayes on Monte Carlo Steroids

<https://papers.nips.cc/paper/6259-variational-bayes-on-monte-carlo-steroids>

**AUTHORS:** Aditya Grover, Stefano Ermon

- HIGHLIGHT: We propose a new class of bounds on the marginal log-likelihood of directed latent variable models.
- 339, TITLE: Showing versus doing: Teaching by demonstration  
<https://papers.nips.cc/paper/6413-showing-versus-doing-teaching-by-demonstration>  
AUTHORS: Mark K. Ho, Michael Littman, James MacGlashan, Fiery Cushman, Joseph L. Austerweil  
HIGHLIGHT: Here, we develop a novel Bayesian model for teaching by demonstration.
- 340, TITLE: Combining Fully Convolutional and Recurrent Neural Networks for 3D Biomedical Image Segmentation  
<https://papers.nips.cc/paper/6448-combining-fully-convolutional-and-recurrent-neural-networks-for-3d-biomedical-image-segmentation>  
AUTHORS: Jianxu Chen, Lin Yang, Yizhe Zhang, Mark Alber, Danny Z. Chen  
HIGHLIGHT: In this paper, we propose a new DL framework for 3D image segmentation, based on a combination of a fully convolutional network (FCN) and a recurrent neural network (RNN), which are responsible for exploiting the intra-slice and inter-slice contexts, respectively.
- 341, TITLE: Maximization of Approximately Submodular Functions  
<https://papers.nips.cc/paper/6236-maximization-of-approximately-submodular-functions>  
AUTHORS: Thibaut Horel, Yaron Singer  
HIGHLIGHT: We study the problem of maximizing a function that is approximately submodular under a cardinality constraint.
- 342, TITLE: A Comprehensive Linear Speedup Analysis for Asynchronous Stochastic Parallel Optimization from Zeroth-Order to First-Order  
<https://papers.nips.cc/paper/6551-a-comprehensive-linear-speedup-analysis-for-asynchronous-stochastic-parallel-optimization-from-zeroth-order-to-first-order>  
AUTHORS: Xiangru Lian, Huan Zhang, Cho-Jui Hsieh, Yijun Huang, Ji Liu  
HIGHLIGHT: This paper provides a comprehensive and generic analysis to study the speedup property for a broad range of asynchronous parallel stochastic algorithms from the zeroth order to the first order methods.
- 343, TITLE: Learning Infinite RBMs with Frank-Wolfe  
<https://papers.nips.cc/paper/6342-learning-infinite-rbms-with-frank-wolfe>  
AUTHORS: Wei Ping, Qiang Liu, Alexander T. Ihler  
HIGHLIGHT: In this work, we propose an infinite restricted Boltzmann machine (RBM), whose maximum likelihood estimation (MLE) corresponds to a constrained convex optimization.
- 344, TITLE: Estimating the Size of a Large Network and its Communities from a Random Sample  
<https://papers.nips.cc/paper/6229-estimating-the-size-of-a-large-network-and-its-communities-from-a-random-sample>  
AUTHORS: Lin Chen, Amin Karbasi, Forrest W. Crawford  
HIGHLIGHT: In this paper we consider a population random graph  $G = (V; E)$  from the stochastic block model (SBM) with  $K$  communities/blocks.
- 345, TITLE: Learning Sensor Multiplexing Design through Back-propagation  
<https://papers.nips.cc/paper/6251-learning-sensor-multiplexing-design-through-back-propagation>  
AUTHORS: Ayan Chakrabarti  
HIGHLIGHT: In this paper, we specifically consider the design and inference problems in a typical color camera---where the sensor is able to measure only one color channel at each pixel location, and computational inference is required to reconstruct a full color image.
- 346, TITLE: On Robustness of Kernel Clustering  
<https://papers.nips.cc/paper/6389-on-robustness-of-kernel-clustering>  
AUTHORS: Bawei Yan, Purnamrita Sarkar  
HIGHLIGHT: On Robustness of Kernel Clustering
- 347, TITLE: High resolution neural connectivity from incomplete tracing data using nonnegative spline regression  
<https://papers.nips.cc/paper/6244-high-resolution-neural-connectivity-from-incomplete-tracing-data-using-nonnegative-spline-regression>  
AUTHORS: Kameron D. Harris, Stefan Mihalas, Eric Shea-Brown  
HIGHLIGHT: We use a novel machine-learning algorithm to meet these challenges and develop a spatially explicit, voxel-scale connectivity map of the mouse visual system.

- 348, TITLE: MoCap-guided Data Augmentation for 3D Pose Estimation in the Wild  
<https://papers.nips.cc/paper/6563-mocap-guided-data-augmentation-for-3d-pose-estimation-in-the-wild>  
AUTHORS: Gregory Rogez, Cordelia Schmid  
HIGHLIGHT: Here, we propose a solution to generate a large set of photorealistic synthetic images of humans with 3D pose annotations.
- 349, TITLE: New Lifiable Classes for First-Order Probabilistic Inference  
<https://papers.nips.cc/paper/6603-new-lifiable-classes-for-first-order-probabilistic-inference>  
AUTHORS: Seyed Mehran Kazemi, Angelika Kimmig, Guy Van den Broeck, David Poole  
HIGHLIGHT: In this paper, we study the domain recursion inference rule, which, despite its central role in early theoretical results on domain-lifted inference, has later been believed redundant.
- 350, TITLE: The Parallel Knowledge Gradient Method for Batch Bayesian Optimization  
<https://papers.nips.cc/paper/6307-the-parallel-knowledge-gradient-method-for-batch-bayesian-optimization>  
AUTHORS: Jian Wu, Peter Frazier  
HIGHLIGHT: In this paper, we develop a novel batch Bayesian optimization algorithm --- the parallel knowledge gradient method.
- 351, TITLE: Improved Regret Bounds for Oracle-Based Adversarial Contextual Bandits  
<https://papers.nips.cc/paper/6400-improved-regret-bounds-for-oracle-based-adversarial-contextual-bandits>  
AUTHORS: Vasilis Syrgkanis, Haipeng Luo, Akshay Krishnamurthy, Robert E. Schapire  
HIGHLIGHT: We propose a new oracle-based algorithm, BISTRO+, for the adversarial contextual bandit problem, where either contexts are drawn i.i.d. or the sequence of contexts is known a priori, but where the losses are picked adversarially.
- 352, TITLE: Consistent Estimation of Functions of Data Missing Non-Monotonically and Not at Random  
<https://papers.nips.cc/paper/6297-consistent-estimation-of-functions-of-data-missing-non-monotonically-and-not-at-random>  
AUTHORS: Ilya Shpitser  
HIGHLIGHT: In this paper, we propose a general class of consistent estimators for cases where data is missing not at random, and missingness status is non-monotonic.
- 353, TITLE: Optimistic Gittins Indices  
<https://papers.nips.cc/paper/6036-optimistic-gittins-indices>  
AUTHORS: Eli Gutin, Vivek Farias  
HIGHLIGHT: The present paper proposes a sequence of 'optimistic' approximations to the Gittins index.
- 354, TITLE: Finite-Dimensional BFRY Priors and Variational Bayesian Inference for Power Law Models  
<https://papers.nips.cc/paper/6348-finite-dimensional-bfry-priors-and-variational-bayesian-inference-for-power-law-models>  
AUTHORS: Juho Lee, Lancelot F. James, Seungjin Choi  
HIGHLIGHT: For illustrative purposes, we describe a simple variational Bayes algorithm for normalized SP mixture models, and demonstrate its usefulness with experiments on synthetic and real-world datasets.
- 355, TITLE: Launch and Iterate: Reducing Prediction Churn  
<https://papers.nips.cc/paper/6053-launch-and-iterate-reducing-prediction-churn>  
AUTHORS: Mahdi Milani Fard, Quentin Cormier, Kevin Canini, Maya Gupta  
HIGHLIGHT: In this paper, we formulate the problem and present a stabilization operator to regularize a classifier towards a previous classifier.
- 356, TITLE: ?Congruent? and ?Opposite? Neurons: Sisters for Multisensory Integration and Segregation  
<https://papers.nips.cc/paper/6317-congruent-and-opposite-neurons-sisters-for-multisensory-integration-and-segregation>  
AUTHORS: Wen-Hao Zhang, He Wang, K. Y. Michael Wong, Si Wu  
HIGHLIGHT: Here, we propose that opposite neurons may serve to encode the disparity information between cues necessary for multisensory segregation.
- 357, TITLE: Learning shape correspondence with anisotropic convolutional neural networks  
<https://papers.nips.cc/paper/6045-learning-shape-correspondence-with-anisotropic-convolutional-neural-networks>  
AUTHORS: Davide Boscaini, Jonathan Masci, Emanuele Rodol?, Michael Bronstein  
HIGHLIGHT: In this paper, we propose Anisotropic Convolutional Neural Network (ACNN), a generalization of classical CNNs to non-Euclidean domains, where classical convolutions are replaced by projections over a set of oriented anisotropic diffusion kernels.

- 358, TITLE: Pairwise Choice Markov Chains  
<https://papers.nips.cc/paper/6287-pairwise-choice-markov-chains>  
AUTHORS: Stephen Ragain, Johan Ugander  
HIGHLIGHT: In this work we introduce the Pairwise Choice Markov Chain (PCMC) model of discrete choice, an inferentially tractable model that does not assume these traditional axioms while still satisfying the foundational axiom of uniform expansion, which can be viewed as a weaker version of Luce's axiom.
- 359, TITLE: NESTT: A Nonconvex Primal-Dual Splitting Method for Distributed and Stochastic Optimization  
<https://papers.nips.cc/paper/6506-nestt-a-nonconvex-primal-dual-splitting-method-for-distributed-and-stochastic-optimization>  
AUTHORS: Davood Hajinezhad, Mingyi Hong, Tuo Zhao, Zhaoran Wang  
HIGHLIGHT: We study a stochastic and distributed algorithm for nonconvex problems whose objective consists a sum of nonconvex smooth functions, plus a nonsmooth regularizer.
- 360, TITLE: Clustering with Same-Cluster Queries  
<https://papers.nips.cc/paper/6449-clustering-with-same-cluster-queries>  
AUTHORS: Hassan Ashtiani, Shrinu Kushagra, Shai Ben-David  
HIGHLIGHT: We propose a framework for Semi-Supervised Active Clustering framework (SSAC), where the learner is allowed to interact with a domain expert, asking whether two given instances belong to the same cluster or not.
- 361, TITLE: Attend, Infer, Repeat: Fast Scene Understanding with Generative Models  
<https://papers.nips.cc/paper/6230-attend-infer-repeat-fast-scene-understanding-with-generative-models>  
AUTHORS: S. M. Ali Eslami, Nicolas Heess, Theophane Weber, Yuval Tassa, David Szepesvari, koray kavukcuoglu, Geoffrey E. Hinton  
HIGHLIGHT: We present a framework for efficient inference in structured image models that explicitly reason about objects.
- 362, TITLE: Parameter Learning for Log-supermodular Distributions  
<https://papers.nips.cc/paper/6402-parameter-learning-for-log-supermodular-distributions>  
AUTHORS: Tatiana Shpakova, Francis Bach  
HIGHLIGHT: In this paper, we focus primarily on parameter estimation in the models from known upper-bounds on the intractable log-partition function.
- 363, TITLE: Deconvolving Feedback Loops in Recommender Systems  
<https://papers.nips.cc/paper/6283-deconvolving-feedback-loops-in-recommender-systems>  
AUTHORS: Ayan Sinha, David F. Gleich, Karthik Ramani  
HIGHLIGHT: We investigate whether it is possible to identify items affected by these feedback loops.
- 364, TITLE: Structured Matrix Recovery via the Generalized Dantzig Selector  
<https://papers.nips.cc/paper/6394-structured-matrix-recovery-via-the-generalized-dantzig-selector>  
AUTHORS: Sheng Chen, Arindam Banerjee  
HIGHLIGHT: In this paper we present non-asymptotic analysis for estimation of generally structured matrices via the generalized Dantzig selector based on sub-Gaussian measurements.
- 365, TITLE: Confusions over Time: An Interpretable Bayesian Model to Characterize Trends in Decision Making  
<https://papers.nips.cc/paper/6234-confusions-over-time-an-interpretable-bayesian-model-to-characterize-trends-in-decision-making>  
AUTHORS: Himabindu Lakkaraju, Jure Leskovec  
HIGHLIGHT: We propose Confusions over Time (CoT), a novel generative framework which facilitates a multi-granular analysis of the decision making process.
- 366, TITLE: Automatic Neuron Detection in Calcium Imaging Data Using Convolutional Networks  
<https://papers.nips.cc/paper/6138-automatic-neuron-detection-in-calcium-imaging-data-using-convolutional-networks>  
AUTHORS: Noah Apthorpe, Alexander Riordan, Robert Aguilar, Jan Homann, Yi Gu, David Tank, H. Sebastian Seung  
HIGHLIGHT: Here we apply a supervised learning approach to this problem and show that convolutional networks can achieve near-human accuracy and superhuman speed.
- 367, TITLE: Designing smoothing functions for improved worst-case competitive ratio in online optimization  
<https://papers.nips.cc/paper/6073-designing-smoothing-functions-for-improved-worst-case-competitive-ratio-in-online-optimization>  
AUTHORS: Reza Eghbali, Maryam Fazel  
HIGHLIGHT: We provide new examples of online problems on the positive orthant and the positive semidefinite cone that satisfy the sufficient condition.

- 368, TITLE: Convergence guarantees for kernel-based quadrature rules in misspecified settings  
<https://papers.nips.cc/paper/6174-convergence-guarantees-for-kernel-based-quadrature-rules-in-misspecified-settings>  
AUTHORS: Motonobu Kanagawa, Bharath K. Sriperumbudur, Kenji Fukumizu  
HIGHLIGHT: Our contribution is in proving that kernel quadratures can be consistent even when the integrand does not belong to the assumed RKHS, i.e., when the integrand is less smooth than assumed.
- 369, TITLE: Unsupervised Learning from Noisy Networks with Applications to Hi-C Data  
<https://papers.nips.cc/paper/6291-unsupervised-learning-from-noisy-networks-with-applications-to-hi-c-data>  
AUTHORS: Bo Wang, Junjie Zhu, Armin Pourshafeie, Oana Ursu, Serafim Batzoglou, Anshul Kundaje  
HIGHLIGHT: In this paper, we propose an optimization framework to mine useful structures from noisy networks in an unsupervised manner.
- 370, TITLE: A Non-generative Framework and Convex Relaxations for Unsupervised Learning  
<https://papers.nips.cc/paper/6533-a-non-generative-framework-and-convex-relaxations-for-unsupervised-learning>  
AUTHORS: Elad Hazan, Tengyu Ma  
HIGHLIGHT: We give a novel formal theoretical framework for unsupervised learning with two distinctive characteristics.
- 371, TITLE: Equality of Opportunity in Supervised Learning  
<https://papers.nips.cc/paper/6374-equality-of-opportunity-in-supervised-learning>  
AUTHORS: Moritz Hardt, Eric Price, ecprice, Nati Srebro  
HIGHLIGHT: We propose a criterion for discrimination against a specified sensitive attribute in supervised learning, where the goal is to predict some target based on available features.
- 372, TITLE: Scaled Least Squares Estimator for GLMs in Large-Scale Problems  
<https://papers.nips.cc/paper/6522-scaled-least-squares-estimator-for-glms-in-large-scale-problems>  
AUTHORS: Murat A. Erdogdu, Lee H. Dicker, Mohsen Bayati  
HIGHLIGHT: We study the problem of efficiently estimating the coefficients of generalized linear models (GLMs) in the large-scale setting where the number of observations  $n$  is much larger than the number of predictors  $p$ , i.e.  $n \gg p$ .
- 373, TITLE: Interpretable Nonlinear Dynamic Modeling of Neural Trajectories  
<https://papers.nips.cc/paper/6543-interpretable-nonlinear-dynamic-modeling-of-neural-trajectories>  
AUTHORS: Yuan Zhao, Il Memming Park  
HIGHLIGHT: We propose a nonlinear time series model aimed at characterizing interpretable dynamics from neural trajectories.
- 374, TITLE: Search Improves Label for Active Learning  
<https://papers.nips.cc/paper/6183-search-improves-label-for-active-learning>  
AUTHORS: Alina Beygelzimer, Daniel J. Hsu, John Langford, Chicheng Zhang  
HIGHLIGHT: We investigate active learning with access to two distinct oracles: LABEL (which is standard) and SEARCH (which is not).
- 375, TITLE: Higher-Order Factorization Machines  
<https://papers.nips.cc/paper/6144-higher-order-factorization-machines>  
AUTHORS: Mathieu Blondel, Akinori Fujino, Naonori Ueda, Masakazu Ishihata  
HIGHLIGHT: In this paper, we present the first generic yet efficient algorithms for training arbitrary-order HOFMs.
- 376, TITLE: Exponential expressivity in deep neural networks through transient chaos  
<https://papers.nips.cc/paper/6322-exponential-expressivity-in-deep-neural-networks-through-transient-chaos>  
AUTHORS: Ben Poole, Subhaneil Lahiri, Maithra Raghu, Jascha Sohl-Dickstein, Surya Ganguli  
HIGHLIGHT: We combine Riemannian geometry with the mean field theory of high dimensional chaos to study the nature of signal propagation in deep neural networks with random weights.
- 377, TITLE: Split LBI: An Iterative Regularization Path with Structural Sparsity  
<https://papers.nips.cc/paper/6288-split-lbi-an-iterative-regularization-path-with-structural-sparsity>  
AUTHORS: Chendi Huang, Xinwei Sun, Jiechao Xiong, Yuan Yao  
HIGHLIGHT: An iterative regularization path with structural sparsity is proposed in this paper based on variable splitting and the Linearized Bregman Iteration, hence called `\emph{Split LBI}`.
- 378, TITLE: An equivalence between high dimensional Bayes optimal inference and M-estimation  
<https://papers.nips.cc/paper/6599-an-equivalence-between-high-dimensional-bayes-optimal-inference-and-m-estimation>

AUTHORS: Madhu Advani, Surya Ganguli  
HIGHLIGHT: In this work we demonstrate how MMSE performance is asymptotically achievable via optimization with an appropriately selected convex penalty and regularization function which are a smoothed version of the widely applied MAP algorithm.

379, TITLE: Synthesizing the preferred inputs for neurons in neural networks via deep generator networks  
<https://papers.nips.cc/paper/6519-synthesizing-the-preferred-inputs-for-neurons-in-neural-networks-via-deep-generator-networks>  
AUTHORS: Anh Nguyen, Alexey Dosovitskiy, Jason Yosinski, Thomas Brox, Jeff Clune  
HIGHLIGHT: Understanding the inner workings of such computational brains is both fascinating basic science that is interesting in its own right---similar to why we study the human brain---and will enable researchers to further improve DNNs.

380, TITLE: Deep Submodular Functions: Definitions and Learning  
<https://papers.nips.cc/paper/6361-deep-submodular-functions-definitions-and-learning>  
AUTHORS: Brian W. Dolhansky, Jeff A. Bilmes  
HIGHLIGHT: We propose and study a new class of submodular functions called deep submodular functions (DSFs).

381, TITLE: Discriminative Gaifman Models  
<https://papers.nips.cc/paper/6098-discriminative-gaifman-models>  
AUTHORS: Mathias Niepert  
HIGHLIGHT: We present discriminative Gaifman models, a novel family of relational machine learning models.

382, TITLE: Leveraging Sparsity for Efficient Submodular Data Summarization  
<https://papers.nips.cc/paper/6382-leveraging-sparsity-for-efficient-submodular-data-summarization>  
AUTHORS: Erik Lindgren, Shanshan Wu, Alexandros G. Dimakis  
HIGHLIGHT: In this paper we establish that these extra assumptions are not necessary---solving the sparsified problem will be almost optimal under the standard assumptions of the problem.

383, TITLE: Local Minimax Complexity of Stochastic Convex Optimization  
<https://papers.nips.cc/paper/6601-local-minimax-complexity-of-stochastic-convex-optimization>  
AUTHORS: sabyasachi chatterjee, John C. Duchi, John Lafferty, Yuancheng Zhu  
HIGHLIGHT: We extend the traditional worst-case, minimax analysis of stochastic convex optimization by introducing a localized form of minimax complexity for individual functions.

384, TITLE: Stochastic Optimization for Large-scale Optimal Transport  
<https://papers.nips.cc/paper/6566-stochastic-optimization-for-large-scale-optimal-transport>  
AUTHORS: Aude Genevay, Marco Cuturi, Gabriel Peyr?, Francis Bach  
HIGHLIGHT: We instantiate these ideas in three different computational setups: (i) when comparing a discrete distribution to another, we show that incremental stochastic optimization schemes can beat the current state of the art finite dimensional OT solver (Sinkhorn's algorithm) ; (ii) when comparing a discrete distribution to a continuous density, a re-formulation (semi-discrete) of the dual program is amenable to averaged stochastic gradient descent, leading to better performance than approximately solving the problem by discretization ; (iii) when dealing with two continuous densities, we propose a stochastic gradient descent over a reproducing kernel Hilbert space (RKHS).

385, TITLE: On Mixtures of Markov Chains  
<https://papers.nips.cc/paper/6078-on-mixtures-of-markov-chains>  
AUTHORS: Rishi Gupta, Ravi Kumar, Sergei Vassilvitskii  
HIGHLIGHT: We study the problem of reconstructing a mixture of Markov chains from the trajectories generated by random walks through the state space.

386, TITLE: Linear Contextual Bandits with Knapsacks  
<https://papers.nips.cc/paper/6292-linear-contextual-bandits-with-knapsacks>  
AUTHORS: Shipra Agrawal, Nikhil Devanur  
HIGHLIGHT: We present algorithms with near-optimal regret bounds for this problem.

387, TITLE: Reconstructing Parameters of Spreading Models from Partial Observations  
<https://papers.nips.cc/paper/6129-reconstructing-parameters-of-spreading-models-from-partial-observations>  
AUTHORS: Andrey Likhov  
HIGHLIGHT: We introduce an efficient dynamic message-passing algorithm, which is able to reconstruct parameters of the spreading model given only partial information on the activation times of nodes in the network.

388, TITLE: Spatiotemporal Residual Networks for Video Action Recognition

- <https://papers.nips.cc/paper/6433-spatiotemporal-residual-networks-for-video-action-recognition>  
AUTHORS: Christoph Feichtenhofer, Axel Pinz, Richard Wildes  
HIGHLIGHT: In this paper, we introduce spatiotemporal ResNets as a combination of these two approaches.
- 389, TITLE: Path-Normalized Optimization of Recurrent Neural Networks with ReLU Activations  
<https://papers.nips.cc/paper/6214-path-normalized-optimization-of-recurrent-neural-networks-with-relu-activations>  
AUTHORS: Behnam Neyshabur, Yuhuai Wu, Ruslan R. Salakhutdinov, Nati Srebro  
HIGHLIGHT: We investigate the parameter-space geometry of recurrent neural networks (RNNs), and develop an adaptation of path-SGD optimization method, attuned to this geometry, that can learn plain RNNs with ReLU activations.
- 390, TITLE: Strategic Attentive Writer for Learning Macro-Actions  
<https://papers.nips.cc/paper/6414-strategic-attentive-writer-for-learning-macro-actions>  
AUTHORS: Alexander Vezhnevets, Volodymyr Mnih, Simon Osindero, Alex Graves, Oriol Vinyals, John Agapiou, koray kavukcuoglu  
HIGHLIGHT: We present a novel deep recurrent neural network architecture that learns to build implicit plans in an end-to-end manner purely by interacting with an environment in reinforcement learning setting.
- 391, TITLE: The Limits of Learning with Missing Data  
<https://papers.nips.cc/paper/6171-the-limits-of-learning-with-missing-data>  
AUTHORS: Brian Bullins, Elad Hazan, Tomer Koren  
HIGHLIGHT: In this well-studied model, we provide the first lower bounds giving a limit on the precision attainable by any algorithm for several variants of regression, notably linear regression with the absolute loss and the squared loss, as well as for classification with the hinge loss.
- 392, TITLE: RETAIN: An Interpretable Predictive Model for Healthcare using Reverse Time Attention Mechanism  
<https://papers.nips.cc/paper/6321-retain-an-interpretable-predictive-model-for-healthcare-using-reverse-time-attention-mechanism>  
AUTHORS: Edward Choi, Mohammad Taha Bahadori, Jimeng Sun, Joshua Kulas, Andy Schuetz, Walter Stewart  
HIGHLIGHT: We addressed this challenge by developing the REverse Time AttentIoN model (RETAIN) for application to Electronic Health Records (EHR) data.
- 393, TITLE: Total Variation Classes Beyond 1d: Minimax Rates, and the Limitations of Linear Smoothers  
<https://papers.nips.cc/paper/6570-total-variation-classes-beyond-1d-minimax-rates-and-the-limitations-of-linear-smoothers>  
AUTHORS: Veeranjaneyulu Sadhanala, Yu-Xiang Wang, Ryan J. Tibshirani  
HIGHLIGHT: We consider the problem of estimating a function defined over  $n$  locations on a  $d$ -dimensional grid (having all side lengths equal to  $n^{1/d}$ ).
- 394, TITLE: Community Detection on Evolving Graphs  
<https://papers.nips.cc/paper/6173-community-detection-on-evolving-graphs>  
AUTHORS: Aris Anagnostopoulos, Jakub Lacki, Silvio Lattanzi, Stefano Leonardi, Mohammad Mahdian  
HIGHLIGHT: In this paper, we study a model of clustering on evolving graphs that captures this aspect of the problem.
- 395, TITLE: Online and Differentially-Private Tensor Decomposition  
<https://papers.nips.cc/paper/6498-online-and-differentially-private-tensor-decomposition>  
AUTHORS: Yining Wang, Anima Anandkumar  
HIGHLIGHT: In this paper, we resolve many of the key algorithmic questions regarding robustness, memory efficiency, and differential privacy of tensor decomposition.
- 396, TITLE: Dimension-Free Iteration Complexity of Finite Sum Optimization Problems  
<https://papers.nips.cc/paper/6087-dimension-free-iteration-complexity-of-finite-sum-optimization-problems>  
AUTHORS: Yossi Arjevani, Ohad Shamir  
HIGHLIGHT: In this work, we extend the framework of Arjevani et al. \cite{arjevani2015lower,arjevani2016iteration} to provide new lower bounds, which are dimension-free, and go beyond the assumptions of current bounds, thereby covering standard finite sum optimization methods, e.g., SAG, SAGA, SVRG, SDCA without duality, as well as stochastic coordinate-descent methods, such as SDCA and accelerated proximal SDCA.
- 397, TITLE: Towards Conceptual Compression  
<https://papers.nips.cc/paper/6542-towards-conceptual-compression>  
AUTHORS: Karol Gregor, Frederic Besse, Danilo Jimenez Rezende, Ivo Danihelka, Daan Wierstra  
HIGHLIGHT: We introduce convolutional DRAW, a homogeneous deep generative model achieving state-of-the-art performance in latent variable image modeling.

- 398, TITLE: Exact Recovery of Hard Thresholding Pursuit  
<https://papers.nips.cc/paper/6432-exact-recovery-of-hard-thresholding-pursuit>  
AUTHORS: Xiaotong Yuan, Ping Li, Tong Zhang  
HIGHLIGHT: In this paper, we bridge this gap by showing, for the first time, that exact recovery of the global sparse minimizer is possible for HTP-style methods under restricted strong condition number bounding conditions.
- 399, TITLE: Data Programming: Creating Large Training Sets, Quickly  
<https://papers.nips.cc/paper/6523-data-programming-creating-large-training-sets-quickly>  
AUTHORS: Alexander J. Ratner, Christopher M. De Sa, Sen Wu, Daniel Selsam, Christopher R?  
HIGHLIGHT: We therefore propose a paradigm for the programmatic creation of training sets called data programming in which users provide a set of labeling functions, which are programs that heuristically label subsets of the data, but that are noisy and may conflict.
- 400, TITLE: Generalization of ERM in Stochastic Convex Optimization: The Dimension Strikes Back  
<https://papers.nips.cc/paper/6467-generalization-of-erm-in-stochastic-convex-optimization-the-dimension-strikes-back>  
AUTHORS: Vitaly Feldman  
HIGHLIGHT: Here we consider the question of how many samples are necessary for ERM to succeed and the closely related question of uniform convergence of  $\mathbb{E} \sum_{k=1}^K \ell_k$  to  $\mathbb{E} \sum_{k=1}^K \ell_k^*$  over  $\mathcal{K}$ .
- 401, TITLE: Dynamic matrix recovery from incomplete observations under an exact low-rank constraint  
<https://papers.nips.cc/paper/6460-dynamic-matrix-recovery-from-incomplete-observations-under-an-exact-low-rank-constraint>  
AUTHORS: Liangbei Xu, Mark Davenport  
HIGHLIGHT: In this paper we aim to address this gap by studying the problem of recovering a dynamically evolving low-rank matrix from incomplete observations.
- 402, TITLE: Fast Distributed Submodular Cover: Public-Private Data Summarization  
<https://papers.nips.cc/paper/6540-fast-distributed-submodular-cover-public-private-data-summarization>  
AUTHORS: Baharan Mirzasoleiman, Morteza Zadimoghaddam, Amin Karbasi  
HIGHLIGHT: In this paper, we introduce the public-private framework of data summarization motivated by privacy concerns in personalized recommender systems and online social services.
- 403, TITLE: Estimating Nonlinear Neural Response Functions using GP Priors and Kronecker Methods  
<https://papers.nips.cc/paper/6153-estimating-nonlinear-neural-response-functions-using-gp-priors-and-kronecker-methods>  
AUTHORS: Cristina Savin, Gasper Tkacik  
HIGHLIGHT: Here we use gaussian process (GP) priors and exploit recent advances in fast GP inference and learning based on Kronecker methods, to efficiently estimate multidimensional nonlinear tuning functions.
- 404, TITLE: Lifelong Learning with Weighted Majority Votes  
<https://papers.nips.cc/paper/6095-lifelong-learning-with-weighted-majority-votes>  
AUTHORS: Anastasia Pentina, Ruth Umer  
HIGHLIGHT: In this work, we consider a setting where the learner encounters a stream of tasks but is able to retain only limited information from each encountered task, such as a learned predictor.
- 405, TITLE: Scaling Memory-Augmented Neural Networks with Sparse Reads and Writes  
<https://papers.nips.cc/paper/6298-scaling-memory-augmented-neural-networks-with-sparse-reads-and-writes>  
AUTHORS: Jack Rae, Jonathan J. Hunt, Ivo Danihelka, Timothy Harley, Andrew W. Senior, Gregory Wayne, Alex Graves, Timothy Lillicrap  
HIGHLIGHT: Here, we present an end-to-end differentiable memory access scheme, which we call Sparse Access Memory (SAM), that retains the representational power of the original approaches whilst training efficiently with very large memories.
- 406, TITLE: Matching Networks for One Shot Learning  
<https://papers.nips.cc/paper/6385-matching-networks-for-one-shot-learning>  
AUTHORS: Oriol Vinyals, Charles Blundell, Timothy Lillicrap, koray kavukcuoglu, Daan Wierstra  
HIGHLIGHT: In this work, we employ ideas from metric learning based on deep neural features and from recent advances that augment neural networks with external memories.
- 407, TITLE: Tight Complexity Bounds for Optimizing Composite Objectives  
<https://papers.nips.cc/paper/6058-tight-complexity-bounds-for-optimizing-composite-objectives>  
AUTHORS: Blake E. Woodworth, Nati Srebro

**HIGHLIGHT:** For non-smooth functions, having access to prox oracles reduces the complexity and we present optimal methods based on smoothing that improve over methods using just gradient accesses.

408, **TITLE:** Graphical Time Warping for Joint Alignment of Multiple Curves  
<https://papers.nips.cc/paper/6269-graphical-time-warping-for-joint-alignment-of-multiple-curves>  
**AUTHORS:** Yizhi Wang, David J. Miller, Kira Poskanzer, Yue Wang, Lin Tian, Guoqiang Yu  
**HIGHLIGHT:** In this paper, we show that the joint alignment problem can be transformed into a network flow problem and thus can be exactly and efficiently solved by the max flow algorithm, with a guarantee of global optimality.

409, **TITLE:** Unsupervised Risk Estimation Using Only Conditional Independence Structure  
<https://papers.nips.cc/paper/6201-unsupervised-risk-estimation-using-only-conditional-independence-structure>  
**AUTHORS:** Jacob Steinhardt, Percy S. Liang  
**HIGHLIGHT:** We show how to estimate a model's test error from unlabeled data, on distributions very different from the training distribution, while assuming only that certain conditional independencies are preserved between train and test.

410, **TITLE:** MetaGrad: Multiple Learning Rates in Online Learning  
<https://papers.nips.cc/paper/6268-metagrad-multiple-learning-rates-in-online-learning>  
**AUTHORS:** Tim van Erven, Wouter M. Koolen  
**HIGHLIGHT:** We present a new method, MetaGrad, that adapts to a much broader class of functions, including exp-concave and strongly convex functions, but also various types of stochastic and non-stochastic functions without any curvature.

411, **TITLE:** Hierarchical Deep Reinforcement Learning: Integrating Temporal Abstraction and Intrinsic Motivation  
<https://papers.nips.cc/paper/6233-hierarchical-deep-reinforcement-learning-integrating-temporal-abstraction-and-intrinsic-motivation>  
**AUTHORS:** Tejas D. Kulkarni, Karthik Narasimhan, Ardavan Saeedi, Josh Tenenbaum  
**HIGHLIGHT:** We present hierarchical-DQN (h-DQN), a framework to integrate hierarchical action-value functions, operating at different temporal scales, with goal-driven intrinsically motivated deep reinforcement learning.

412, **TITLE:** High Dimensional Structured Superposition Models  
<https://papers.nips.cc/paper/6079-high-dimensional-structured-superposition-models>  
**AUTHORS:** Qilong Gu, Arindam Banerjee  
**HIGHLIGHT:** In this paper, we consider general superposition models which allow sum of any number of component parameters, and each component structure can be characterized by any norm.

413, **TITLE:** Joint quantile regression in vector-valued RKHSs  
<https://papers.nips.cc/paper/6239-joint-quantile-regression-in-vector-valued-rkhs>  
**AUTHORS:** Maxime Sangnier, Olivier Fercoq, Florence d'Alché-Buc  
**HIGHLIGHT:** Joint quantile regression in vector-valued RKHSs

414, **TITLE:** The Forget-me-not Process  
<https://papers.nips.cc/paper/6055-the-forget-me-not-process>  
**AUTHORS:** Kieran Milan, Joel Veness, James Kirkpatrick, Michael Bowling, Anna Koop, Demis Hassabis  
**HIGHLIGHT:** We introduce the Forget-me-not Process, an efficient, non-parametric meta-algorithm for online probabilistic sequence prediction for piecewise stationary, repeating sources.

415, **TITLE:** Wasserstein Training of Restricted Boltzmann Machines  
<https://papers.nips.cc/paper/6248-wasserstein-training-of-restricted-boltzmann-machines>  
**AUTHORS:** Grégoire Montavon, Klaus-Robert Müller, Marco Cuturi  
**HIGHLIGHT:** We propose in this work a novel approach for Boltzmann machine training which assumes that a meaningful metric between observations is known.

416, **TITLE:** Communication-Optimal Distributed Clustering  
<https://papers.nips.cc/paper/6562-communication-optimal-distributed-clustering>  
**AUTHORS:** Jiecao Chen, He Sun, David Woodruff, Qin Zhang  
**HIGHLIGHT:** In this work, we study both graph and geometric clustering problems in two distributed models: (1) a point-to-point model, and (2) a model with a broadcast channel.

417, **TITLE:** Probing the Compositionality of Intuitive Functions  
<https://papers.nips.cc/paper/6130-probing-the-compositionality-of-intuitive-functions>  
**AUTHORS:** Eric Schulz, Josh Tenenbaum, David K. Duvenaud, Maarten Speekenbrink, Samuel J. Gershman

**HIGHLIGHT:** Taking inspiration from other areas of cognitive science, we propose that this is accomplished by harnessing compositionality: complex structure is decomposed into simpler building blocks.

418, **TITLE:** Ladder Variational Autoencoders  
<https://papers.nips.cc/paper/6275-ladder-variational-autoencoders>  
**AUTHORS:** Casper Kaae Sønderby, Tapani Raiko, Lars Maaløe, Søren Kaae Sønderby, Ole Winther  
**HIGHLIGHT:** We propose a new inference model, the Ladder Variational Autoencoder, that recursively corrects the generative distribution by a data dependent approximate likelihood in a process resembling the recently proposed Ladder Network.

419, **TITLE:** The Multiple Quantile Graphical Model  
<https://papers.nips.cc/paper/6092-the-multiple-quantile-graphical-model>  
**AUTHORS:** Alnur Ali, J. Zico Kolter, Ryan J. Tibshirani  
**HIGHLIGHT:** We introduce the Multiple Quantile Graphical Model (MQGM), which extends the neighborhood selection approach of Meinshausen and Buhlmann for learning sparse graphical models.

420, **TITLE:** Threshold Learning for Optimal Decision Making  
<https://papers.nips.cc/paper/6494-threshold-learning-for-optimal-decision-making>  
**AUTHORS:** Nathan F. Lepora  
**HIGHLIGHT:** We examine threshold learning by constructing a reward function that averages over many trials to Wald's cost function that defines decision optimality.

421, **TITLE:** Unsupervised Feature Extraction by Time-Contrastive Learning and Nonlinear ICA  
<https://papers.nips.cc/paper/6395-unsupervised-feature-extraction-by-time-contrastive-learning-and-nonlinear-ica>  
**AUTHORS:** Aapo Hyvarinen, Hiroshi Morioka  
**HIGHLIGHT:** Nonlinear independent component analysis (ICA) provides an appealing framework for unsupervised feature learning, but the models proposed so far are not identifiable.

422, **TITLE:** Can Active Memory Replace Attention?  
<https://papers.nips.cc/paper/6295-can-active-memory-replace-attention>  
**AUTHORS:** Lukasz Kaiser, Samy Bengio  
**HIGHLIGHT:** We analyze this shortcoming in this paper and propose an extended model of active memory that matches existing attention models on neural machine translation and generalizes better to longer sentences.

423, **TITLE:** Minimax Optimal Alternating Minimization for Kernel Nonparametric Tensor Learning  
<https://papers.nips.cc/paper/6419-minimax-optimal-alternating-minimization-for-kernel-nonparametric-tensor-learning>  
**AUTHORS:** Taiji Suzuki, Heishiro Kanagawa, Hayato Kobayashi, Nobuyuki Shimizu, Yukihiro Tagami  
**HIGHLIGHT:** In this paper, we show that the alternating minimization method achieves linear convergence as an optimization algorithm and that the generalization error of the resultant estimator yields the minimax optimality.

424, **TITLE:** The Product Cut  
<https://papers.nips.cc/paper/6226-the-product-cut>  
**AUTHORS:** Thomas Laurent, James von Brecht, Xavier Bresson, arthur szlam  
**HIGHLIGHT:** We introduce a theoretical and algorithmic framework for multi-way graph partitioning that relies on a multiplicative cut-based objective.

425, **TITLE:** Learning Sparse Gaussian Graphical Models with Overlapping Blocks  
<https://papers.nips.cc/paper/6097-learning-sparse-gaussian-graphical-models-with-overlapping-blocks>  
**AUTHORS:** Mohammad Javad Hosseini, Su-In Lee  
**HIGHLIGHT:** We present a novel framework, called GRAB (GRaphical models with overlApping Blocks), to capture densely connected components in a network estimate.

426, **TITLE:** Yggdrasil: An Optimized System for Training Deep Decision Trees at Scale  
<https://papers.nips.cc/paper/6366-yggdrasil-an-optimized-system-for-training-deep-decision-trees-at-scale>  
**AUTHORS:** Firas Abuzaid, Joseph K. Bradley, Feynman T. Liang, Andrew Feng, Lee Yang, Matei Zaharia, Ameet S. Talwalkar  
**HIGHLIGHT:** We present Yggdrasil, a new distributed tree learning method that outperforms existing methods by up to 24x.

427, **TITLE:** Average-case hardness of RIP certification  
<https://papers.nips.cc/paper/6132-average-case-hardness-of-rip-certification>  
**AUTHORS:** Tengyao Wang, Quentin Berthet, Yaniv Plan

**HIGHLIGHT:** We consider the framework of average-case certifiers, that never wrongly declare that a matrix is RIP, while being often correct for random instances.

428, **TITLE:** A forward model at Purkinje cell synapses facilitates cerebellar anticipatory control  
<https://papers.nips.cc/paper/6151-a-forward-model-at-purkinje-cell-synapses-facilitates-cerebellar-anticipatory-control>

**AUTHORS:** Ivan Herreros, Xerxes Arsiwalla, Paul Verschure

**HIGHLIGHT:** Based on that interaction we design the counterfactual predictive control (CFPC) architecture, an anticipatory adaptive motor control scheme in which a feed-forward module, based on the cerebellum, steers an error feedback controller with counterfactual error signals.

429, **TITLE:** Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering  
<https://papers.nips.cc/paper/6081-convolutional-neural-networks-on-graphs-with-fast-localized-spectral-filtering>

**AUTHORS:** Michaël Defferrard, Xavier Bresson, Pierre Vandergheynst

**HIGHLIGHT:** In this work, we are interested in generalizing convolutional neural networks (CNNs) from low-dimensional regular grids, where image, video and speech are represented, to high-dimensional irregular domains, such as social networks, brain connectomes or words' embedding, represented by graphs.

430, **TITLE:** CliqueCNN: Deep Unsupervised Exemplar Learning

<https://papers.nips.cc/paper/6387-cliquecnn-deep-unsupervised-exemplar-learning>

**AUTHORS:** Miguel A. Bautista, Artsiom Sanakoyeu, Ekaterina Tikhoncheva, Bjorn Ommer

**HIGHLIGHT:** Given weak estimates of local distance we propose a single optimization problem to extract batches of samples with mutually consistent relations.

431, **TITLE:** Large-Scale Price Optimization via Network Flow

<https://papers.nips.cc/paper/6301-large-scale-price-optimization-via-network-flow>

**AUTHORS:** Shinji Ito, Ryohei Fujimaki

**HIGHLIGHT:** To cope with this problem, this paper proposes a novel approach based on network flow algorithms.

432, **TITLE:** Online Pricing with Strategic and Patient Buyers

<https://papers.nips.cc/paper/6415-online-pricing-with-strategic-and-patient-buyers>

**AUTHORS:** Michal Feldman, Tomer Koren, Roi Livni, Yishay Mansour, Aviv Zohar

**HIGHLIGHT:** In this setting, we give an algorithm that attains  $O(T^{2/3})$  regret over any sequence of  $T$  buyers with respect to the best fixed price in hindsight, and prove that no algorithm can perform better in the worst case.

433, **TITLE:** Global Optimality of Local Search for Low Rank Matrix Recovery

<https://papers.nips.cc/paper/6271-global-optimality-of-local-search-for-low-rank-matrix-recovery>

**AUTHORS:** Srinadh Bhojanapalli, Behnam Neyshabur, Nati Srebro

**HIGHLIGHT:** We show that there are no spurious local minima in the non-convex factorized parametrization of low-rank matrix recovery from incoherent linear measurements.

434, **TITLE:** Phased LSTM: Accelerating Recurrent Network Training for Long or Event-based Sequences

<https://papers.nips.cc/paper/6310-phased-lstm-accelerating-recurrent-network-training-for-long-or-event-based-sequences>

**AUTHORS:** Daniel Neil, Michael Pfeiffer, Shih-Chii Liu

**HIGHLIGHT:** In this work, we introduce the Phased LSTM model, which extends the LSTM unit by adding a new time gate.

435, **TITLE:** Improving PAC Exploration Using the Median Of Means

<https://papers.nips.cc/paper/6577-improving-pac-exploration-using-the-median-of-means>

**AUTHORS:** Jason Pazis, Ronald E. Parr, Jonathan P. How

**HIGHLIGHT:** We present the first application of the median of means in a PAC exploration algorithm for MDPs.

436, **TITLE:** Infinite Hidden Semi-Markov Modulated Interaction Point Process

<https://papers.nips.cc/paper/6243-infinite-hidden-semi-markov-modulated-interaction-point-process>

**AUTHORS:** matt zhang, Peng Lin, Peng Lin, Ting Guo, Yang Wang, Yang Wang, Fang Chen

**HIGHLIGHT:** In this paper, we propose a Bayesian nonparametric approach that considers both types of correlations via unifying and generalizing hidden semi-Markov model and interaction point process model.

437, **TITLE:** Cooperative Inverse Reinforcement Learning

<https://papers.nips.cc/paper/6420-cooperative-inverse-reinforcement-learning>

**AUTHORS:** Dylan Hadfield-Menell, Stuart J. Russell, Pieter Abbeel, Anca Dragan

HIGHLIGHT: We propose a formal definition of the value alignment problem as cooperative inverse reinforcement learning (CIRL).

438, TITLE: Spatio-Temporal Hilbert Maps for Continuous Occupancy Representation in Dynamic Environments  
<https://papers.nips.cc/paper/6541-spatio-temporal-hilbert-maps-for-continuous-occupancy-representation-in-dynamic-environments>  
AUTHORS: Ransalu Senanayake, Lionel Ott, Simon O'Callaghan, Fabio T. Ramos  
HIGHLIGHT: We propose a method to propagate motion uncertainty into the kernel using a hierarchical model.

439, TITLE: Select-and-Sample for Spike-and-Slab Sparse Coding  
<https://papers.nips.cc/paper/6276-select-and-sample-for-spike-and-slab-sparse-coding>  
AUTHORS: Abdul-Saboor Sheikh, J?rg L?cke  
HIGHLIGHT: Here, we study an approach that can efficiently be scaled while maintaining a richly structured posterior approximation under these conditions.

440, TITLE: Tractable Operations for Arithmetic Circuits of Probabilistic Models  
<https://papers.nips.cc/paper/6363-tractable-operations-for-arithmetic-circuits-of-probabilistic-models>  
AUTHORS: Yujia Shen, Arthur Choi, Adnan Darwiche  
HIGHLIGHT: We consider tractable representations of probability distributions and the polytime operations they support.

441, TITLE: Greedy Feature Construction  
<https://papers.nips.cc/paper/6557-greedy-feature-construction>  
AUTHORS: Dino Oglic, Thomas G?rtner  
HIGHLIGHT: We present an effective method for supervised feature construction.

442, TITLE: Mistake Bounds for Binary Matrix Completion  
<https://papers.nips.cc/paper/6567-mistake-bounds-for-binary-matrix-completion>  
AUTHORS: Mark Herbster, Stephen Pasteris, Massimiliano Pontil  
HIGHLIGHT: We propose a Matrix Exponentiated Gradient algorithm [1] to solve this problem.

443, TITLE: Data driven estimation of Laplace-Beltrami operator  
<https://papers.nips.cc/paper/6210-data-driven-estimation-of-laplace-beltrami-operator>  
AUTHORS: Frederic Chazal, Ilaria Giulini, Bertrand Michel  
HIGHLIGHT: In this paper, we address this problem for the unnormalized graph Laplacian by establishing an oracle inequality that opens the door to a well-founded data-driven procedure for the bandwidth selection.

444, TITLE: Tracking the Best Expert in Non-stationary Stochastic Environments  
<https://papers.nips.cc/paper/6536-tracking-the-best-expert-in-non-stationary-stochastic-environments>  
AUTHORS: Chen-Yu Wei, Yi-Te Hong, Chi-Jen Lu  
HIGHLIGHT: We not only propose algorithms with upper bound guarantee, but prove their matching lower bounds as well.

445, TITLE: Learning to learn by gradient descent by gradient descent  
<https://papers.nips.cc/paper/6461-learning-to-learn-by-gradient-descent-by-gradient-descent>  
AUTHORS: Marcin Andrychowicz, Misha Denil, Sergio Gomez, Matthew W. Hoffman, David Pfau, Tom Schaul, Brendan Shillingford, Nando de Freitas  
HIGHLIGHT: In this paper we show how the design of an optimization algorithm can be cast as a learning problem, allowing the algorithm to learn to exploit structure in the problems of interest in an automatic way.

446, TITLE: Kernel Observers: Systems-Theoretic Modeling and Inference of Spatiotemporally Evolving Processes  
<https://papers.nips.cc/paper/6189-kernel-observers-systems-theoretic-modeling-and-inference-of-spatiotemporally-evolving-processes>  
AUTHORS: Hassan A. Kingravi, Harshal R. Maske, Girish Chowdhary  
HIGHLIGHT: We consider the problem of estimating the latent state of a spatiotemporally evolving continuous function using very few sensor measurements.

447, TITLE: Quantum Perceptron Models  
<https://papers.nips.cc/paper/6401-quantum-perceptron-models>  
AUTHORS: Ashish Kapoor, Nathan Wiebe, Krysta Svore  
HIGHLIGHT: We develop two quantum algorithms for perceptron learning.

448, TITLE: Guided Policy Search via Approximate Mirror Descent

- <https://papers.nips.cc/paper/6105-guided-policy-search-via-approximate-mirror-descent>  
AUTHORS: William H. Montgomery, Sergey Levine  
HIGHLIGHT: We derive a new guided policy search algorithm that is simpler and provides appealing improvement and convergence guarantees in simplified convex and linear settings, and show that in the more general nonlinear setting, the error in the projection step can be bounded.
- 449, TITLE: The Power of Optimization from Samples  
<https://papers.nips.cc/paper/6447-the-power-of-optimization-from-samples>  
AUTHORS: Eric Balkanski, Aviad Rubinstein, Yaron Singer  
HIGHLIGHT: In this paper we show that for any monotone submodular function with curvature  $c$  there is a  $(1 - c)/(1 + c - c^2)$  approximation algorithm for maximization under cardinality constraints when polynomially-many samples are drawn from the uniform distribution over feasible sets.
- 450, TITLE: Deep Exploration via Bootstrapped DQN  
<https://papers.nips.cc/paper/6501-deep-exploration-via-bootstrapped-dqn>  
AUTHORS: Ian Osband, Charles Blundell, Alexander Pritzel, Benjamin Van Roy  
HIGHLIGHT: We demonstrate that bootstrapped DQN can combine deep exploration with deep neural networks for exponentially faster learning than any dithering strategy.
- 451, TITLE: A Multi-step Inertial Forward-Backward Splitting Method for Non-convex Optimization  
<https://papers.nips.cc/paper/6285-a-multi-step-inertial-forward-backward-splitting-method-for-non-convex-optimization>  
AUTHORS: Jingwei Liang, Jalal Fadili, Gabriel Peyr?  
HIGHLIGHT: In this paper, we propose a multi-step inertial Forward-Backward splitting algorithm for minimizing the sum of two non-necessarily convex functions, one of which is proper lower semi-continuous while the other is differentiable with a Lipschitz continuous gradient.
- 452, TITLE: Scaling Factorial Hidden Markov Models: Stochastic Variational Inference without Messages  
<https://papers.nips.cc/paper/6534-scaling-factorial-hidden-markov-models-stochastic-variational-inference-without-messages>  
AUTHORS: Yin Cheng Ng, Pawel M. Chilinski, Ricardo Silva  
HIGHLIGHT: We propose a scalable inference and learning algorithm for FHMMs that draws on ideas from the stochastic variational inference, neural network and copula literatures.
- 453, TITLE: Convolutional Neural Fabrics  
<https://papers.nips.cc/paper/6304-convolutional-neural-fabrics>  
AUTHORS: Shreyas Saxena, Jakob Verbeek  
HIGHLIGHT: Instead of aiming to select a single optimal architecture, we propose a "fabric" that embeds an exponentially large number of architectures.
- 454, TITLE: Adaptive Newton Method for Empirical Risk Minimization to Statistical Accuracy  
<https://papers.nips.cc/paper/6262-adaptive-newton-method-for-empirical-risk-minimization-to-statistical-accuracy>  
AUTHORS: Aryan Mokhtari, Hadi Daneshmand, Aurelien Lucchi, Thomas Hofmann, Alejandro Ribeiro  
HIGHLIGHT: We introduce Ada Newton as an adaptive algorithm that uses Newton's method with adaptive sample sizes.
- 455, TITLE: A Sparse Interactive Model for Matrix Completion with Side Information  
<https://papers.nips.cc/paper/6265-a-sparse-interactive-model-for-matrix-completion-with-side-information>  
AUTHORS: Jin Lu, Guannan Liang, Jiangwen Sun, Jinbo Bi  
HIGHLIGHT: We propose a novel sparse formulation that explicitly models the interaction between the row and column side features to approximate the matrix entries.
- 456, TITLE: Coresets for Scalable Bayesian Logistic Regression  
<https://papers.nips.cc/paper/6486-coresets-for-scalable-bayesian-logistic-regression>  
AUTHORS: Jonathan Huggins, Trevor Campbell, Tamara Broderick  
HIGHLIGHT: In this paper, we develop an efficient coreset construction algorithm for Bayesian logistic regression models.
- 457, TITLE: Agnostic Estimation for Misspecified Phase Retrieval Models  
<https://papers.nips.cc/paper/6094-agnostic-estimation-for-misspecified-phase-retrieval-models>  
AUTHORS: Matey Neykov, Zhaoran Wang, Han Liu  
HIGHLIGHT: Based on this model, we propose a significant semi-parametric generalization called misspecified phase retrieval (MPR), in which  $Y = f(X^{\top} \beta^* + \epsilon)$  with unknown  $f$  and  $\text{Cov}(Y, (X^{\top} \beta^*)^2) \geq 0$ .

- 458, TITLE: Linear Relaxations for Finding Diverse Elements in Metric Spaces  
<https://papers.nips.cc/paper/6500-linear-relaxations-for-finding-diverse-elements-in-metric-spaces>  
AUTHORS: Aditya Bhaskara, Mehrdad Ghadiri, Vahab Mirrokni, Ola Svensson  
HIGHLIGHT: The goal of this paper is to develop a novel linear programming (LP) framework that allows us to design approximation algorithms for such problems.
- 459, TITLE: Binarized Neural Networks  
<https://papers.nips.cc/paper/6573-binarized-neural-networks>  
AUTHORS: Itay Hubara, Matthieu Courbariaux, Daniel Soudry, Ran El-Yaniv, Yoshua Bengio  
HIGHLIGHT: We introduce a method to train Binarized Neural Networks (BNNs) - neural networks with binary weights and activations at run-time.
- 460, TITLE: Local Maxima in the Likelihood of Gaussian Mixture Models: Structural Results and Algorithmic Consequences  
<https://papers.nips.cc/paper/6324-local-maxima-in-the-likelihood-of-gaussian-mixture-models-structural-results-and-algorithmic-consequences>  
AUTHORS: Chi Jin, Yuchen Zhang, Sivaraman Balakrishnan, Martin J. Wainwright, Michael I. Jordan  
HIGHLIGHT: We provide two fundamental results on the population (infinite-sample) likelihood function of Gaussian mixture models with  $M \geq 3$  components.
- 461, TITLE: Memory-Efficient Backpropagation Through Time  
<https://papers.nips.cc/paper/6221-memory-efficient-backpropagation-through-time>  
AUTHORS: Audrunas Gruslys, Remi Munos, Ivo Danihelka, Marc Lanctot, Alex Graves  
HIGHLIGHT: We propose a novel approach to reduce memory consumption of the backpropagation through time (BPTT) algorithm when training recurrent neural networks (RNNs).
- 462, TITLE: Bayesian Optimization with Robust Bayesian Neural Networks  
<https://papers.nips.cc/paper/6117-bayesian-optimization-with-robust-bayesian-neural-networks>  
AUTHORS: Jost Tobias Springenberg, Aaron Klein, Stefan Falkner, Frank Hutter  
HIGHLIGHT: We present a general approach for using flexible parametric models (neural networks) for Bayesian optimization, staying as close to a truly Bayesian treatment as possible.
- 463, TITLE: Learnable Visual Markers  
<https://papers.nips.cc/paper/6323-learnable-visual-markers>  
AUTHORS: Oleg Grinchuk, Vadim Lebedev, Victor Lempitsky  
HIGHLIGHT: We propose a new approach to designing visual markers (analogous to QR-codes, markers for augmented reality, and robotic fiducial tags) based on the advances in deep generative networks.
- 464, TITLE: Fast Algorithms for Robust PCA via Gradient Descent  
<https://papers.nips.cc/paper/6445-fast-algorithms-for-robust-pca-via-gradient-descent>  
AUTHORS: Xinyang Yi, Dohyung Park, Yudong Chen, Constantine Caramanis  
HIGHLIGHT: This paper presents and analyzes a non-convex optimization approach that greatly reduces the computational complexity of the above problems, compared to the best available algorithms.
- 465, TITLE: One-vs-Each Approximation to Softmax for Scalable Estimation of Probabilities  
<https://papers.nips.cc/paper/6468-one-vs-each-approximation-to-softmax-for-scalable-estimation-of-probabilities>  
AUTHORS: Michalis Titsias  
HIGHLIGHT: Here, we introduce an efficient approximation to softmax probabilities which takes the form of a rigorous lower bound on the exact probability.
- 466, TITLE: Learning Deep Embeddings with Histogram Loss  
<https://papers.nips.cc/paper/6464-learning-deep-embeddings-with-histogram-loss>  
AUTHORS: Evgeniya Ustinova, Victor Lempitsky  
HIGHLIGHT: We suggest a new loss for learning deep embeddings.
- 467, TITLE: Spectral Learning of Dynamic Systems from Nonequilibrium Data  
<https://papers.nips.cc/paper/6191-spectral-learning-of-dynamic-systems-from-nonequilibrium-data>  
AUTHORS: Hao Wu, Frank Noe

**HIGHLIGHT:** In this paper, we investigate the properties of spectral learning without this assumption due to the requirements of analyzing large-time scale systems, and show that the equilibrium dynamics of a system can be extracted from nonequilibrium observation data by imposing an equilibrium constraint.

468, **TITLE:** Fast Mixing Markov Chains for Strongly Rayleigh Measures, DPPs, and Constrained Sampling  
<https://papers.nips.cc/paper/6182-fast-mixing-markov-chains-for-strongly-rayleigh-measures-dpps-and-constrained-sampling>  
**AUTHORS:** Chengtao Li, Suvrit Sra, Stefanie Jegelka  
**HIGHLIGHT:** Our first main result is for MCMC sampling from Strongly Rayleigh (SR) measures, for which we present sharp polynomial bounds on the mixing time.

469, **TITLE:** Mapping Estimation for Discrete Optimal Transport  
<https://papers.nips.cc/paper/6312-mapping-estimation-for-discrete-optimal-transport>  
**AUTHORS:** Micha?l Perrot, Nicolas Courty, R?mi Flamary, Amaury Habrard  
**HIGHLIGHT:** In this paper we propose a new way to jointly learn the coupling and an approximation of the transport map.

470, **TITLE:** Batched Gaussian Process Bandit Optimization via Determinantal Point Processes  
<https://papers.nips.cc/paper/6452-batched-gaussian-process-bandit-optimization-via-determinantal-point-processes>  
**AUTHORS:** Tarun Kathuria, Amit Deshpande, Pushmeet Kohli  
**HIGHLIGHT:** In this paper, we propose a new approach for parallelizing Bayesian optimization by modeling the diversity of a batch via Determinantal point processes (DPPs) whose kernels are learned automatically.

471, **TITLE:** Protein contact prediction from amino acid co-evolution using convolutional networks for graph-valued images  
<https://papers.nips.cc/paper/6488-protein-contact-prediction-from-amino-acid-co-evolution-using-convolutional-networks-for-graph-valued-images>  
**AUTHORS:** Vladimir Golkov, Marcin J. Skwark, Antonij Golkov, Alexey Dosovitskiy, Thomas Brox, Jens Meiler, Daniel Cremers  
**HIGHLIGHT:** We propose several methods for treating such "graph-valued images" in a convolutional network.

472, **TITLE:** Linear Feature Encoding for Reinforcement Learning  
<https://papers.nips.cc/paper/6305-linear-feature-encoding-for-reinforcement-learning>  
**AUTHORS:** Zhao Song, Ronald E. Parr, Xuejun Liao, Lawrence Carin  
**HIGHLIGHT:** This paper develops and evaluates a theory of linear feature encoding.

473, **TITLE:** A Minimax Approach to Supervised Learning  
<https://papers.nips.cc/paper/6247-a-minimax-approach-to-supervised-learning>  
**AUTHORS:** Farzan Farnia, David Tse  
**HIGHLIGHT:** In this paper, we address this question by introducing a generalization of the maximum entropy principle.

474, **TITLE:** Edge-exchangeable graphs and sparsity  
<https://papers.nips.cc/paper/6586-edge-exchangeable-graphs-and-sparsity>  
**AUTHORS:** Diana Cai, Trevor Campbell, Tamara Broderick  
**HIGHLIGHT:** We present an alternative notion of exchangeability for random graphs, which we call edge exchangeability, in which the distribution of a graph sequence is invariant to the order of the edges.

475, **TITLE:** A Locally Adaptive Normal Distribution  
<https://papers.nips.cc/paper/6503-a-locally-adaptive-normal-distribution>  
**AUTHORS:** Georgios Arvanitidis, Lars K. Hansen, S?ren Hauberg  
**HIGHLIGHT:** We develop a maximum likelihood algorithm to infer the distribution parameters that relies on a combination of gradient descent and Monte Carlo integration.

476, **TITLE:** Completely random measures for modelling block-structured sparse networks  
<https://papers.nips.cc/paper/6521-completely-random-measures-for-modelling-block-structured-sparse-networks>  
**AUTHORS:** Tue Herlau, Mikkel N. Schmidt, Morten M?rup  
**HIGHLIGHT:** In this work we re-introduce the use of block-structure for network models obeying Kallenberg's notion of exchangeability and thereby obtain a collapsed model which both admits the inference of block-structure and edge inhomogeneity.

477, **TITLE:** Sparse Support Recovery with Non-smooth Loss Functions  
<https://papers.nips.cc/paper/6559-sparse-support-recovery-with-non-smooth-loss-functions>  
**AUTHORS:** K?vin Degraux, Gabriel Peyr?, Jalal Fadili, Laurent Jacques

HIGHLIGHT: In this paper, we study the support recovery guarantees of underdetermined sparse regression using the  $\ell_1$ -norm as a regularizer and a non-smooth loss function for data fidelity.

478, TITLE: Neurons Equipped with Intrinsic Plasticity Learn Stimulus Intensity Statistics  
<https://papers.nips.cc/paper/6582-neurons-equipped-with-intrinsic-plasticity-learn-stimulus-intensity-statistics>  
AUTHORS: Travis Monk, Cristina Savin, J?rg L?cke  
HIGHLIGHT: We introduce a novel generative mixture model that accounts for the class-specific statistics of stimulus intensities, and we derive a neural circuit that learns the input classes and their intensities.

479, TITLE: Learning values across many orders of magnitude  
<https://papers.nips.cc/paper/6076-learning-values-across-many-orders-of-magnitude>  
AUTHORS: Hado P. van Hasselt, Arthur Guez, Arthur Guez, Matteo Hessel, Volodymyr Mnih, David Silver  
HIGHLIGHT: We propose to adaptively normalize the targets used in the learning updates.

480, TITLE: Adaptive Smoothed Online Multi-Task Learning  
<https://papers.nips.cc/paper/6434-adaptive-smoothed-online-multi-task-learning>  
AUTHORS: Keerthiram Murugesan, Hanxiao Liu, Jaime Carbonell, Yiming Yang  
HIGHLIGHT: This paper addresses the challenge of jointly learning both the per-task model parameters and the inter-task relationships in a multi-task online learning setting.

481, TITLE: Safe Exploration in Finite Markov Decision Processes with Gaussian Processes  
<https://papers.nips.cc/paper/6358-safe-exploration-in-finite-markov-decision-processes-with-gaussian-processes>  
AUTHORS: Matteo Turchetta, Felix Berkenkamp, Andreas Krause  
HIGHLIGHT: In this paper, we address the problem of safely exploring finite Markov decision processes (MDP).

482, TITLE: Probabilistic Linear Multistep Methods  
<https://papers.nips.cc/paper/6356-probabilistic-linear-multistep-methods>  
AUTHORS: Onur Teymur, Kostas Zygalakis, Ben Calderhead  
HIGHLIGHT: We present a derivation and theoretical investigation of the Adams-Bashforth and Adams-Moulton family of linear multistep methods for solving ordinary differential equations, starting from a Gaussian process (GP) framework.

483, TITLE: Stochastic Three-Composite Convex Minimization  
<https://papers.nips.cc/paper/6127-stochastic-three-composite-convex-minimization>  
AUTHORS: Alp Yurtsever, Bang Cong Vu, Volkan Cevher  
HIGHLIGHT: We propose a stochastic optimization method for the minimization of the sum of three convex functions, one of which has Lipschitz continuous gradient as well as restricted strong convexity.

484, TITLE: Using Fast Weights to Attend to the Recent Past  
<https://papers.nips.cc/paper/6057-using-fast-weights-to-attend-to-the-recent-past>  
AUTHORS: Jimmy Ba, Geoffrey E. Hinton, Volodymyr Mnih, Joel Z. Leibo, Catalin Ionescu  
HIGHLIGHT: By using fast weights we can avoid the need to store copies of neural activity patterns.

485, TITLE: Maximal Sparsity with Deep Networks?  
<https://papers.nips.cc/paper/6346-maximal-sparsity-with-deep-networks>  
AUTHORS: Bo Xin, Yizhou Wang, Wen Gao, David Wipf, Baoyuan Wang  
HIGHLIGHT: In contrast, we demonstrate both theoretically and empirically the potential for a trained deep network to recover minimal  $\ell_0$ -norm representations in regimes where existing methods fail.

486, TITLE: Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings  
<https://papers.nips.cc/paper/6228-man-is-to-computer-programmer-as-woman-is-to-homemaker-debiasing-word-embeddings>  
AUTHORS: Tolga Bolukbasi, Kai-Wei Chang, James Y. Zou, Venkatesh Saligrama, Adam T. Kalai  
HIGHLIGHT: This raises concerns because their widespread use, as we describe, often tends to amplify these biases.

487, TITLE: beta-risk: a New Surrogate Risk for Learning from Weakly Labeled Data  
<https://papers.nips.cc/paper/6156-beta-risk-a-new-surrogate-risk-for-learning-from-weakly-labeled-data>  
AUTHORS: Valentina Zantedeschi, R?mi Emonet, Marc Sebban  
HIGHLIGHT: This paper presents a generic framework to deal with these weakly labeled scenarios.

- 488, TITLE: Learning Additive Exponential Family Graphical Models via  $\ell_{2,1}$ -norm Regularized M-Estimation  
[https://papers.nips.cc/paper/6106-learning-additive-exponential-family-graphical-models-via-ell\\_21-norm-regularized-m-estimation](https://papers.nips.cc/paper/6106-learning-additive-exponential-family-graphical-models-via-ell_21-norm-regularized-m-estimation)  
AUTHORS: Xiaotong Yuan, Ping Li, Tong Zhang, Qingshan Liu, Guangcan Liu  
HIGHLIGHT: We propose two  $\ell_{2,1}$ -norm regularized maximum likelihood estimators to learn the model parameters from i.i.d. samples.
- 489, TITLE: Backprop KF: Learning Discriminative Deterministic State Estimators  
<https://papers.nips.cc/paper/6090-backprop-kf-learning-discriminative-deterministic-state-estimators>  
AUTHORS: Tuomas Haarnoja, Anurag Ajay, Sergey Levine, Pieter Abbeel  
HIGHLIGHT: We present an alternative approach where the parameters of the latent state distribution are directly optimized as a deterministic computation graph, resulting in a simple and effective gradient descent algorithm for training discriminative state estimators.
- 490, TITLE: LightRNN: Memory and Computation-Efficient Recurrent Neural Networks  
<https://papers.nips.cc/paper/6512-lightmn-memory-and-computation-efficient-recurrent-neural-networks>  
AUTHORS: Xiang Li, Tao Qin, Jian Yang, Tie-Yan Liu  
HIGHLIGHT: In this work, we propose a novel technique to tackle this challenge.
- 491, TITLE: Fast recovery from a union of subspaces  
<https://papers.nips.cc/paper/6484-fast-recovery-from-a-union-of-subspaces>  
AUTHORS: Chinmay Hegde, Piotr Indyk, Ludwig Schmidt  
HIGHLIGHT: We address the problem of recovering a high-dimensional but structured vector from linear observations in a general setting where the vector can come from an arbitrary union of subspaces.
- 492, TITLE: Incremental Variational Sparse Gaussian Process Regression  
<https://papers.nips.cc/paper/6473-incremental-variational-sparse-gaussian-process-regression>  
AUTHORS: Ching-An Cheng, Byron Boots  
HIGHLIGHT: We propose a novel incremental learning algorithm for variational sparse GPR based on stochastic mirror ascent of probability densities in reproducing kernel Hilbert space.
- 493, TITLE: A Consistent Regularization Approach for Structured Prediction  
<https://papers.nips.cc/paper/6093-a-consistent-regularization-approach-for-structured-prediction>  
AUTHORS: Carlo Ciliberto, Lorenzo Rosasco, Alessandro Rudi  
HIGHLIGHT: We propose and analyze a regularization approach for structured prediction problems.
- 494, TITLE: Clustering Signed Networks with the Geometric Mean of Laplacians  
<https://papers.nips.cc/paper/6164-clustering-signed-networks-with-the-geometric-mean-of-laplacians>  
AUTHORS: Pedro Mercado, Francesco Tudisco, Matthias Hein  
HIGHLIGHT: As a solution we propose to use the geometric mean of the Laplacians of positive and negative part and show that it outperforms the existing approaches.
- 495, TITLE: An urn model for majority voting in classification ensembles  
<https://papers.nips.cc/paper/6120-an-urn-model-for-majority-voting-in-classification-ensembles>  
AUTHORS: Victor Soto, Alberto Suarez, Gonzalo Martinez-Muoz  
HIGHLIGHT: In this work we analyze the class prediction of parallel randomized ensembles by majority voting as an urn model.
- 496, TITLE: Avoiding Imposters and Delinquents: Adversarial Crowdsourcing and Peer Prediction  
<https://papers.nips.cc/paper/6440-avoiding-imposters-and-delinquents-adversarial-crowdsourcing-and-peer-prediction>  
AUTHORS: Jacob Steinhardt, Gregory Valiant, Moses Charikar  
HIGHLIGHT: We consider a crowdsourcing model in which  $n$  workers are asked to rate the quality of  $n$  items previously generated by other workers.
- 497, TITLE: Fast and accurate spike sorting of high-channel count probes with KiloSort  
<https://papers.nips.cc/paper/6326-fast-and-accurate-spike-sorting-of-high-channel-count-probes-with-kilosort>  
AUTHORS: Marius Pachitariu, Nicholas A. Steinmetz, Shabnam N. Kadir, Matteo Carandini, Kenneth D. Harris  
HIGHLIGHT: Here we introduce KiloSort, a new integrated spike sorting framework that uses template matching both during spike detection and during spike clustering.

- 498, TITLE: Combining Adversarial Guarantees and Stochastic Fast Rates in Online Learning  
<https://papers.nips.cc/paper/6474-combining-adversarial-guarantees-and-stochastic-fast-rates-in-online-learning>  
AUTHORS: Wouter M. Koolen, Peter Grunwald, Tim van Erven  
HIGHLIGHT: We consider online learning algorithms that guarantee worst-case regret rates in adversarial environments (so they can be deployed safely and will perform robustly), yet adapt optimally to favorable stochastic environments (so they will perform well in a variety of settings of practical importance).
- 499, TITLE: Ancestral Causal Inference  
<https://papers.nips.cc/paper/6266-ancestral-causal-inference>  
AUTHORS: Sara Magliacane, Tom Claassen, Joris M. Mooij  
HIGHLIGHT: We present a novel method that reduces the combinatorial explosion of the search space by using a more coarse-grained representation of causal information, drastically reducing computation time.
- 500, TITLE: More Supervision, Less Computation: Statistical-Computational Tradeoffs in Weakly Supervised Learning  
<https://papers.nips.cc/paper/6518-more-supervision-less-computation-statistical-computational-tradeoffs-in-weakly-supervised-learning>  
AUTHORS: Xinyang Yi, Zhaoran Wang, Zhuoran Yang, Constantine Caramanis, Han Liu  
HIGHLIGHT: In this paper, we characterize the effect of  $\alpha$  by establishing the information-theoretic and computational boundaries, namely, the minimax-optimal statistical accuracy that can be achieved by all algorithms, and polynomial-time algorithms under an oracle computational model.
- 501, TITLE: Tagger: Deep Unsupervised Perceptual Grouping  
<https://papers.nips.cc/paper/6067-tagger-deep-unsupervised-perceptual-grouping>  
AUTHORS: Klaus Greff, Antti Rasmus, Mathias Berglund, Tele Hao, Harri Valpola, Jrgen Schmidhuber  
HIGHLIGHT: We present a framework for efficient perceptual inference that explicitly reasons about the segmentation of its inputs and features.
- 502, TITLE: An Efficient Streaming Algorithm for the Submodular Cover Problem  
<https://papers.nips.cc/paper/6175-an-efficient-streaming-algorithm-for-the-submodular-cover-problem>  
AUTHORS: Ashkan Norouzi-Fard, Abbas Bazzi, Ilija Bogunovic, Marwa El Halabi, Ya-Ping Hsieh, Volkan Cevher  
HIGHLIGHT: We initiate the study of the classical Submodular Cover (SC) problem in the data streaming model which we refer to as the Streaming Submodular Cover (SSC).
- 503, TITLE: Interaction Networks for Learning about Objects, Relations and Physics  
<https://papers.nips.cc/paper/6418-interaction-networks-for-learning-about-objects-relations-and-physics>  
AUTHORS: Peter Battaglia, Razvan Pascanu, Matthew Lai, Danilo Jimenez Rezende, koray kavukcuoglu  
HIGHLIGHT: Here we introduce the interaction network, a model which can reason about how objects in complex systems interact, supporting dynamical predictions, as well as inferences about the abstract properties of the system.
- 504, TITLE: Efficient state-space modularization for planning: theory, behavioral and neural signatures  
<https://papers.nips.cc/paper/6320-efficient-state-space-modularization-for-planning-theory-behavioral-and-neural-signatures>  
AUTHORS: Daniel McNamee, Daniel M. Wolpert, Mate Lengyel  
HIGHLIGHT: Here, we develop a normative theory of efficient state-space representations which partitions an environment into distinct modules by minimizing the average (information theoretic) description length of planning within the environment, thereby optimally trading off the complexity of planning across and within modules.
- 505, TITLE: Provable Efficient Online Matrix Completion via Non-convex Stochastic Gradient Descent  
<https://papers.nips.cc/paper/6204-provable-efficient-online-matrix-completion-via-non-convex-stochastic-gradient-descent>  
AUTHORS: Chi Jin, Sham M. Kakade, Praneeth Netrapalli  
HIGHLIGHT: In this paper, we propose the first provable, efficient online algorithm for matrix completion.
- 506, TITLE: Online Bayesian Moment Matching for Topic Modeling with Unknown Number of Topics  
<https://papers.nips.cc/paper/6077-online-bayesian-moment-matching-for-topic-modeling-with-unknown-number-of-topics>  
AUTHORS: Wei-Shou Hsu, Pascal Poupart  
HIGHLIGHT: We propose two new models that extend LDA in a simple and intuitive fashion by directly expressing a distribution over the number of topics.
- 507, TITLE: Computing and maximizing influence in linear threshold and triggering models  
<https://papers.nips.cc/paper/6347-computing-and-maximizing-influence-in-linear-threshold-and-triggering-models>  
AUTHORS: Justin T. Khim, Varun Jog, Po-Ling Loh  
HIGHLIGHT: We establish upper and lower bounds for the influence of a set of nodes in certain types of contagion models.

- 508, TITLE: Coevolutionary Latent Feature Processes for Continuous-Time User-Item Interactions  
<https://papers.nips.cc/paper/6480-coevolutionary-latent-feature-processes-for-continuous-time-user-item-interactions>  
AUTHORS: Yichen Wang, Nan Du, Rakshit Trivedi, Le Song  
HIGHLIGHT: We propose a coevolutionary latent feature process model that accurately captures the coevolving nature of users' and items' feature.
- 509, TITLE: Optimal Learning for Multi-pass Stochastic Gradient Methods  
<https://papers.nips.cc/paper/6213-optimal-learning-for-multi-pass-stochastic-gradient-methods>  
AUTHORS: Junhong Lin, Lorenzo Rosasco  
HIGHLIGHT: We analyze the learning properties of the stochastic gradient method when multiple passes over the data and mini-batches are allowed.
- 510, TITLE: Generative Adversarial Imitation Learning  
<https://papers.nips.cc/paper/6391-generative-adversarial-imitation-learning>  
AUTHORS: Jonathan Ho, Stefano Ermon  
HIGHLIGHT: We propose a new general framework for directly extracting a policy from data as if it were obtained by reinforcement learning following inverse reinforcement learning.
- 511, TITLE: Latent Attention For If-Then Program Synthesis  
<https://papers.nips.cc/paper/6284-latent-attention-for-if-then-program-synthesis>  
AUTHORS: Chang Liu, Xinyun Chen, Richard Shin, Mingcheng Chen, Dawn Song  
HIGHLIGHT: In this work, we consider a simple yet important sub-problem: translation from textual descriptions to If-Then programs.
- 512, TITLE: Dual Space Gradient Descent for Online Learning  
<https://papers.nips.cc/paper/6560-dual-space-gradient-descent-for-online-learning>  
AUTHORS: Trung Le, Tu Nguyen, Vu Nguyen, Dinh Phung  
HIGHLIGHT: To address all of these aforementioned challenges, we present in this paper the Dual Space Gradient Descent (DualSGD), a novel framework that utilizes random features as an auxiliary space to maintain information from data points removed during budget maintenance.
- 513, TITLE: Riemannian SVRG: Fast Stochastic Optimization on Riemannian Manifolds  
<https://papers.nips.cc/paper/6515-riemannian-svrg-fast-stochastic-optimization-on-riemannian-manifolds>  
AUTHORS: Hongyi Zhang, Sashank J. Reddi, Suvrit Sra  
HIGHLIGHT: We introduce  $\text{Riemannian SVRG}$ , a new variance reduced Riemannian optimization method.
- 514, TITLE: Professor Forcing: A New Algorithm for Training Recurrent Networks  
<https://papers.nips.cc/paper/6099-professor-forcing-a-new-algorithm-for-training-recurrent-networks>  
AUTHORS: Alex M. Lamb, Anirudh Goyal ALIAS PARTH GOYAL, Ying Zhang, Saizheng Zhang, Aaron C. Courville, Yoshua Bengio  
HIGHLIGHT: We introduce the Professor Forcing algorithm, which uses adversarial domain adaptation to encourage the dynamics of the recurrent network to be the same when training the network and when sampling from the network over multiple time steps.
- 515, TITLE: Learning brain regions via large-scale online structured sparse dictionary learning  
<https://papers.nips.cc/paper/6352-learning-brain-regions-via-large-scale-online-structured-sparse-dictionary-learning>  
AUTHORS: Elvis DOHMATOB, Arthur Mensch, Gael Varoquaux, Bertrand Thirion  
HIGHLIGHT: We propose a multivariate online dictionary-learning method for obtaining decompositions of brain images with structured and sparse components (aka atoms).
- 516, TITLE: Efficient Neural Codes under Metabolic Constraints  
<https://papers.nips.cc/paper/6470-efficient-neural-codes-under-metabolic-constraints>  
AUTHORS: Zhuo Wang, Xue-Xin Wei, Alan A. Stocker, Daniel D. Lee  
HIGHLIGHT: Here we formulate a coding framework which explicitly deals with noise and the metabolic costs associated with the neural representation of information, and analytically derive the optimal neural code for monotonic response functions and arbitrary stimulus distributions.
- 517, TITLE: Approximate maximum entropy principles via Goemans-Williamson with applications to provable variational methods

- <https://papers.nips.cc/paper/6169-approximate-maximum-entropy-principles-via-goemans-williamson-with-applications-to-provable-variational-methods>  
AUTHORS: Andrej Risteski, Yuanzhi Li  
HIGHLIGHT: We provide computationally efficient versions of this principle when the mean parameters are pairwise moments: we design distributions that approximately match given pairwise moments, while having entropy which is comparable to the maximum entropy distribution matching those moments.
- 518, TITLE: Efficient High-Order Interaction-Aware Feature Selection Based on Conditional Mutual Information  
<https://papers.nips.cc/paper/6584-efficient-high-order-interaction-aware-feature-selection-based-on-conditional-mutual-information>  
AUTHORS: Alexander Shishkin, Anastasia Bezzubtseva, Alexey Drutsa, Iliia Shishkov, Ekaterina Gladkikh, Gleb Gusev, Pavel Serdyukov  
HIGHLIGHT: This study introduces a novel feature selection approach CMICOT, which is a further evolution of filter methods with sequential forward selection (SFS) whose scoring functions are based on conditional mutual information (MI).
- 519, TITLE: Bayesian Intermittent Demand Forecasting for Large Inventories  
<https://papers.nips.cc/paper/6313-bayesian-intermittent-demand-forecasting-for-large-inventories>  
AUTHORS: Matthias W. Seeger, David Salinas, Valentin Flunkert  
HIGHLIGHT: We present a scalable and robust Bayesian method for demand forecasting in the context of a large e-commerce platform, paying special attention to intermittent and bursty target statistics.
- 520, TITLE: Visual Question Answering with Question Representation Update (QRU)  
<https://papers.nips.cc/paper/6261-visual-question-answering-with-question-representation-update-qr>  
AUTHORS: Ruiyu Li, Jiaya Jia  
HIGHLIGHT: Our method aims at reasoning over natural language questions and visual images.
- 521, TITLE: Learning Parametric Sparse Models for Image Super-Resolution  
<https://papers.nips.cc/paper/6378-learning-parametric-sparse-models-for-image-super-resolution>  
AUTHORS: Yongbo Li, Weisheng Dong, Xuemei Xie, GUANGMING Shi, Xin Li, Donglai Xu  
HIGHLIGHT: In this paper, we propose to combine those two lines of ideas for image super-resolution.
- 522, TITLE: Blazing the trails before beating the path: Sample-efficient Monte-Carlo planning  
<https://papers.nips.cc/paper/6253-blazing-the-trails-before-beating-the-path-sample-efficient-monte-carlo-planning>  
AUTHORS: Jean-Bastien Grill, Michal Valko, Remi Munos  
HIGHLIGHT: We propose a new algorithm, TrailBlazer, able to handle MDPs with a finite or an infinite number of transitions from state-action to next states.
- 523, TITLE: Asynchronous Parallel Greedy Coordinate Descent  
<https://papers.nips.cc/paper/6070-asynchronous-parallel-greedy-coordinate-descent>  
AUTHORS: Yang You, Xiangru Lian, Ji Liu, Hsiang-Fu Yu, Inderjit S. Dhillon, James Demmel, Cho-Jui Hsieh  
HIGHLIGHT: In this paper, we propose and study an Asynchronous parallel Greedy Coordinate Descent (Asy-GCD) algorithm for minimizing a smooth function with bounded constraints.
- 524, TITLE: Iterative Refinement of the Approximate Posterior for Directed Belief Networks  
<https://papers.nips.cc/paper/6311-iterative-refinement-of-the-approximate-posterior-for-directed-belief-networks>  
AUTHORS: Devon Hjelm, Ruslan R. Salakhutdinov, Kyunghyun Cho, Nebojsa Jojic, Vince Calhoun, Junyoung Chung  
HIGHLIGHT: To address these issues, we introduce an iterative refinement procedure for improving the approximate posterior of the recognition network and show that training with the refined posterior is competitive with state-of-the-art methods.
- 525, TITLE: Assortment Optimization Under the Mallows model  
<https://papers.nips.cc/paper/6224-assortment-optimization-under-the-mallows-model>  
AUTHORS: Antoine Desir, Vineet Goyal, Srikanth Jagabathula, Danny Segev  
HIGHLIGHT: Our key contributions are an efficiently computable closed-form expression for the choice probability under the Mallows model and a compact mixed integer linear program (MIP) formulation for the assortment problem.
- 526, TITLE: Disease Trajectory Maps  
<https://papers.nips.cc/paper/6177-disease-trajectory-maps>  
AUTHORS: Peter Schulam, Raman Arora  
HIGHLIGHT: In this paper, we focus on answering two questions that can be asked using these databases of longitudinal EHR data.

- 527, TITLE: Multistage Campaigning in Social Networks  
<https://papers.nips.cc/paper/6102-multistage-campaigning-in-social-networks>  
AUTHORS: Mehrdad Farajtabar, Xiaojing Ye, Sahar Harati, Le Song, Hongyuan Zha  
HIGHLIGHT: We consider control problems for multi-stage campaigning over social networks.
- 528, TITLE: Learning in Games: Robustness of Fast Convergence  
<https://papers.nips.cc/paper/6133-learning-in-games-robustness-of-fast-convergence>  
AUTHORS: Dylan J. Foster, Zhiyuan Li, Thodoris Lykouris, Karthik Sridharan, Eva Tardos  
HIGHLIGHT: In the bandit setting we present a new algorithm which provides a "small loss"-type bound with improved dependence on the number of actions in utility settings, and is both simple and efficient.
- 529, TITLE: Improved Variational Inference with Inverse Autoregressive Flow  
<https://papers.nips.cc/paper/6581-improved-variational-inference-with-inverse-autoregressive-flow>  
AUTHORS: Durk P. Kingma, Tim Salimans, Rafal Jozefowicz, Xi Chen, Ilya Sutskever, Max Welling  
HIGHLIGHT: We propose a new type of normalizing flow, inverse autoregressive flow (IAF), that, in contrast to earlier published flows, scales well to high-dimensional latent spaces.
- 530, TITLE: Algorithms and matching lower bounds for approximately-convex optimization  
<https://papers.nips.cc/paper/6576-algorithms-and-matching-lower-bounds-for-approximately-convex-optimization>  
AUTHORS: Andrej Risteski, Yuanzhi Li  
HIGHLIGHT: In this paper, we strengthen the known  $\epsilon$  lower bounds on the trade-off between  $\epsilon$  and  $\epsilon$  substantially, and exhibit an algorithm that matches these lower bounds for a large class of convex bodies.
- 531, TITLE: Unified Methods for Exploiting Piecewise Linear Structure in Convex Optimization  
<https://papers.nips.cc/paper/6043-unified-methods-for-exploiting-piecewise-linear-structure-in-convex-optimization>  
AUTHORS: Tyler B. Johnson, Carlos Guestrin  
HIGHLIGHT: By considering a novel problem formulation—the minimization of a sum of piecewise functions—we describe a principled and general mechanism for exploiting piecewise linear structure in convex optimization.
- 532, TITLE: Kernel Bayesian Inference with Posterior Regularization  
<https://papers.nips.cc/paper/6235-kernel-bayesian-inference-with-posterior-regularization>  
AUTHORS: Yang Song, Jun Zhu, Yong Ren  
HIGHLIGHT: We propose a vector-valued regression problem whose solution is equivalent to the reproducing kernel Hilbert space (RKHS) embedding of the Bayesian posterior distribution.
- 533, TITLE: Neural Universal Discrete Denoiser  
<https://papers.nips.cc/paper/6497-neural-universal-discrete-denoiser>  
AUTHORS: Taesup Moon, Seonwoo Min, Byunghan Lee, Sungroh Yoon  
HIGHLIGHT: We present a new framework of applying deep neural networks (DNN) to devise a universal discrete denoiser.
- 534, TITLE: Optimal Architectures in a Solvable Model of Deep Networks  
<https://papers.nips.cc/paper/6330-optimal-architectures-in-a-solvable-model-of-deep-networks>  
AUTHORS: Jonathan Kadmon, Haim Sompolinsky  
HIGHLIGHT: The purpose of this work is to advance our theoretical understanding of the computational benefits of these architectures.
- 535, TITLE: Conditional Image Generation with PixelCNN Decoders  
<https://papers.nips.cc/paper/6527-conditional-image-generation-with-pixelcnn-decoders>  
AUTHORS: Aaron van den Oord, Nal Kalchbrenner, Lasse Espeholt, koray kavukcuoglu, Oriol Vinyals, Alex Graves  
HIGHLIGHT: This work explores conditional image generation with a new image density model based on the PixelCNN architecture.
- 536, TITLE: Supervised Learning with Tensor Networks  
<https://papers.nips.cc/paper/6211-supervised-learning-with-tensor-networks>  
AUTHORS: Edwin Stoudenmire, David J. Schwab  
HIGHLIGHT: We discuss an interpretation of the additional structure imparted by the tensor network to the learned model.
- 537, TITLE: Multi-step learning and underlying structure in statistical models  
<https://papers.nips.cc/paper/6197-multi-step-learning-and-underlying-structure-in-statistical-models>  
AUTHORS: Maia Fraser

HIGHLIGHT: A related principle arises in transfer-learning where Baxter (2000) proposed a theoretical framework to study how learning multiple tasks transforms the inductive bias of a learner.

538, TITLE: Structure-Blind Signal Recovery  
<https://papers.nips.cc/paper/6063-structure-blind-signal-recovery>  
AUTHORS: Dmitry Ostrovsky, Zaid Harchaoui, Anatoli Juditsky, Arkadi S. Nemirovski  
HIGHLIGHT: We propose a new family of estimators to recover signals observed in Gaussian noise.

539, TITLE: An Architecture for Deep, Hierarchical Generative Models  
<https://papers.nips.cc/paper/6141-an-architecture-for-deep-hierarchical-generative-models>  
AUTHORS: Philip Bachman  
HIGHLIGHT: We present an architecture which lets us train deep, directed generative models with many layers of latent variables.  
We include deterministic paths between all latent variables and the generated output, and provide a richer set of connections between computations for inference and generation, which enables more effective communication of information throughout the model during training.

540, TITLE: Feature selection in functional data classification with recursive maxima hunting  
<https://papers.nips.cc/paper/6392-feature-selection-in-functional-data-classification-with-recursive-maxima-hunting>  
AUTHORS: Jos? L. Torrecilla, Alberto Su?rez  
HIGHLIGHT: In this work, we introduce recursive maxima hunting (RMH) for variable selection in classification problems with functional data.

541, TITLE: Achieving budget-optimality with adaptive schemes in crowdsourcing  
<https://papers.nips.cc/paper/6124-achieving-budget-optimality-with-adaptive-schemes-in-crowdsourcing>  
AUTHORS: Ashish Khetan, Sewoong Oh  
HIGHLIGHT: Under this generalized Dawid-Skene model, we characterize the fundamental trade-off between budget and accuracy, and introduce a novel adaptive scheme that matches this fundamental limit.

542, TITLE: Near-Optimal Smoothing of Structured Conditional Probability Matrices  
<https://papers.nips.cc/paper/6199-near-optimal-smoothing-of-structured-conditional-probability-matrices>  
AUTHORS: Moein Falahatgar, Mesrob I. Ohannessian, Alon Orlitsky  
HIGHLIGHT: We derive an iterative algorithm that extends classical non-negative matrix factorization to naturally incorporate additive smoothing and prove that it converges to the stationary points of a penalized empirical risk.

543, TITLE: Supervised Word Mover's Distance  
<https://papers.nips.cc/paper/6139-supervised-word-movers-distance>  
AUTHORS: Gao Huang, Chuan Guo, Matt J. Kusner, Yu Sun, Fei Sha, Kilian Q. Weinberger  
HIGHLIGHT: In this paper we propose an efficient technique to learn a supervised metric, which we call the Supervised WMD (S-WMD) metric.

544, TITLE: Exploiting Tradeoffs for Exact Recovery in Heterogeneous Stochastic Block Models  
<https://papers.nips.cc/paper/6574-exploiting-tradeoffs-for-exact-recovery-in-heterogeneous-stochastic-block-models>  
AUTHORS: Amin Jalali, Qiyang Han, Ioana Dumitriu, Maryam Fazel  
HIGHLIGHT: In this paper, we consider the SBM in its full generality, where there is no restriction on the number and sizes of communities or how they grow with the number of nodes, as well as on the connectivity probabilities inside or across communities.

545, TITLE: Full-Capacity Unitary Recurrent Neural Networks  
<https://papers.nips.cc/paper/6327-full-capacity-unitary-recurrent-neural-networks>  
AUTHORS: Scott Wisdom, Thomas Powers, John Hershey, Jonathan Le Roux, Les Atlas  
HIGHLIGHT: To address this question, we propose full-capacity uRNNs that optimize their recurrence matrix over all unitary matrices, leading to significantly improved performance over uRNNs that use a restricted-capacity recurrence matrix.

546, TITLE: Threshold Bandits, With and Without Censored Feedback  
<https://papers.nips.cc/paper/6149-threshold-bandits-with-and-without-censored-feedback>  
AUTHORS: Jacob D. Abernethy, Kareem Amin, Ruihao Zhu  
HIGHLIGHT: We consider the `\emph{Threshold Bandit}` setting, a variant of the classical multi-armed bandit problem in which the reward on each round depends on a piece of side information known as a `\emph{threshold value}`.

547, TITLE: Understanding the Effective Receptive Field in Deep Convolutional Neural Networks

- <https://papers.nips.cc/paper/6203-understanding-the-effective-receptive-field-in-deep-convolutional-neural-networks>  
AUTHORS: Wenjie Luo, Yujia Li, Raquel Urtasun, Richard Zemel  
HIGHLIGHT: We introduce the notion of an effective receptive field size, and show that it both has a Gaussian distribution and only occupies a fraction of the full theoretical receptive field size.
- 548, TITLE: Learning Supervised PageRank with Gradient-Based and Gradient-Free Optimization Methods  
<https://papers.nips.cc/paper/6565-learning-supervised-pagerank-with-gradient-based-and-gradient-free-optimization-methods>  
AUTHORS: Lev Bogolubsky, Pavel Dvurechenskii, Alexander Gasnikov, Gleb Gusev, Yurii Nesterov, Andrei M. Raigorodskii, Aleksey Tikhonov, Maksim Zhukovskii  
HIGHLIGHT: In this paper, we consider a non-convex loss-minimization problem of learning Supervised PageRank models, which can account for features of nodes and edges.
- 549, TITLE:  $k^*$ -Nearest Neighbors: From Global to Local  
<https://papers.nips.cc/paper/6373-k-nearest-neighbors-from-global-to-local>  
AUTHORS: Oren Anava, Kfir Levy  
HIGHLIGHT: In this paper we offer a simple approach to locally weighted regression/classification, where we make the bias-variance tradeoff explicit.
- 550, TITLE: Normalized Spectral Map Synchronization  
<https://papers.nips.cc/paper/6128-normalized-spectral-map-synchronization>  
AUTHORS: Yanyao Shen, Qixing Huang, Nati Srebro, Sujay Sanghavi  
HIGHLIGHT: In this paper, we provide theoretical justifications for spectral techniques for the map synchronization problem, i.e., it takes as input a collection of objects and noisy maps estimated between pairs of objects, and outputs clean maps between all pairs of objects.
- 551, TITLE: Beyond Exchangeability: The Chinese Voting Process  
<https://papers.nips.cc/paper/6362-beyond-exchangeability-the-chinese-voting-process>  
AUTHORS: Moontae Lee, Seok Hyun Jin, David Mimno  
HIGHLIGHT: We propose the Chinese Voting Process (CVP) which models the evolution of helpfulness votes as a self-reinforcing process dependent on position and presentation biases.
- 552, TITLE: A posteriori error bounds for joint matrix decomposition problems  
<https://papers.nips.cc/paper/6424-a-posteriori-error-bounds-for-joint-matrix-decomposition-problems>  
AUTHORS: Nicolo Colombo, Nikos Vlassis  
HIGHLIGHT: We consider the problem of approximate joint matrix triangularization when the matrices in  $M$  are jointly diagonalizable and real, but we only observe a set  $M'$  of noise perturbed versions of the matrices in  $M$ .
- 553, TITLE: A Bayesian method for reducing bias in neural representational similarity analysis  
<https://papers.nips.cc/paper/6131-a-bayesian-method-for-reducing-bias-in-neural-representational-similarity-analysis>  
AUTHORS: Ming Bo Cai, Nicolas W. Schuck, Jonathan W. Pillow, Yael Niv  
HIGHLIGHT: We propose an alternative Bayesian framework for computing representational similarity in which we treat the covariance structure of neural activity patterns as a hyper-parameter in a generative model of the neural data, and directly estimate this covariance structure from imaging data while marginalizing over the unknown activity patterns.
- 554, TITLE: Online ICA: Understanding Global Dynamics of Nonconvex Optimization via Diffusion Processes  
<https://papers.nips.cc/paper/6306-online-ica-understanding-global-dynamics-of-nonconvex-optimization-via-diffusion-processes>  
AUTHORS: Chris Junchi Li, Zhaoran Wang, Han Liu  
HIGHLIGHT: In this paper, we propose a new analytic paradigm based on diffusion processes to characterize the global dynamics of nonconvex statistical optimization.
- 555, TITLE: Following the Leader and Fast Rates in Linear Prediction: Curved Constraint Sets and Other Regularities  
<https://papers.nips.cc/paper/6455-following-the-leader-and-fast-rates-in-linear-prediction-curved-constraint-sets-and-other-regularities>  
AUTHORS: Ruitong Huang, Tor Lattimore, Andr?s Gy?rgy, Csaba Szepesvari  
HIGHLIGHT: In this paper we ask whether there are other "lucky" settings when FTL achieves sublinear, "small" regret.
- 556, TITLE: SDP Relaxation with Randomized Rounding for Energy Disaggregation  
<https://papers.nips.cc/paper/6555-sdp-relaxation-with-randomized-rounding-for-energy-disaggregation>  
AUTHORS: Kiarash Shaloudegi, Andr?s Gy?rgy, Csaba Szepesvari, Wilsun Xu  
HIGHLIGHT: We develop a scalable, computationally efficient method for the task of energy disaggregation for home appliance monitoring.

- 557, TITLE: Recovery Guarantee of Non-negative Matrix Factorization via Alternating Updates  
<https://papers.nips.cc/paper/6417-recovery-guarantee-of-non-negative-matrix-factorization-via-alternating-updates>  
AUTHORS: Yuanzhi Li, Yingyu Liang, Andrej Risteski  
HIGHLIGHT: This paper proposes an algorithm that alternates between decoding the weights and updating the features, and shows that assuming a generative model of the data, it provably recovers the ground-truth under fairly mild conditions.
- 558, TITLE: Unsupervised Learning of 3D Structure from Images  
<https://papers.nips.cc/paper/6600-unsupervised-learning-of-3d-structure-from-images>  
AUTHORS: Danilo Jimenez Rezende, S. M. Ali Eslami, Shakir Mohamed, Peter Battaglia, Max Jaderberg, Nicolas Heess  
HIGHLIGHT: In this paper we learn strong deep generative models of 3D structures, and recover these structures from 2D images via probabilistic inference.
- 559, TITLE: Poisson-Gamma dynamical systems  
<https://papers.nips.cc/paper/6083-poisson-gamma-dynamical-systems>  
AUTHORS: Aaron Schein, Hanna Wallach, Mingyuan Zhou  
HIGHLIGHT: This paper presents a dynamical system based on the Poisson-Gamma construction for sequentially observed multivariate count data.
- 560, TITLE: Gaussian Processes for Survival Analysis  
<https://papers.nips.cc/paper/6443-gaussian-processes-for-survival-analysis>  
AUTHORS: Tamara Fernandez, Nicolas Rivera, Yee Whye Teh  
HIGHLIGHT: We introduce a semi-parametric Bayesian model for survival analysis.
- 561, TITLE: Dual Decomposed Learning with Factorwise Oracle for Structural SVM of Large Output Domain  
<https://papers.nips.cc/paper/6422-dual-decomposed-learning-with-factorwise-oracle-for-structural-svm-of-large-output-domain>  
AUTHORS: Ian En-Hsu Yen, Xiangru Huang, Kai Zhong, Ruohan Zhang, Pradeep K. Ravikumar, Inderjit S. Dhillon  
HIGHLIGHT: In this work, we show that, by decomposing training of Structural Support Vector Machine (SVM) into a series of multiclass SVM problems connected through messages, one can replace expensive structured oracle with Factorwise Maximization Oracle (FMO) that allows efficient implementation of complexity sublinear to the factor domain.
- 562, TITLE: Optimal Binary Classifier Aggregation for General Losses  
<https://papers.nips.cc/paper/6597-optimal-binary-classifier-aggregation-for-general-losses>  
AUTHORS: Akshay Balsubramani, Yoav S. Freund  
HIGHLIGHT: We address the problem of aggregating an ensemble of predictors with known loss bounds in a semi-supervised binary classification setting, to minimize prediction loss incurred on the unlabeled data.
- 563, TITLE: Disentangling factors of variation in deep representation using adversarial training  
<https://papers.nips.cc/paper/6051-disentangling-factors-of-variation-in-deep-representation-using-adversarial-training>  
AUTHORS: Michael F. Mathieu, Junbo Jake Zhao, Junbo Zhao, Aditya Ramesh, Pablo Sprechmann, Yann LeCun  
HIGHLIGHT: We propose a deep generative model for learning to distill the hidden factors of variation within a set of labeled observations into two complementary codes.
- 564, TITLE: A primal-dual method for conic constrained distributed optimization problems  
<https://papers.nips.cc/paper/6242-a-primal-dual-method-for-conic-constrained-distributed-optimization-problems>  
AUTHORS: Necdet Serhat Aybat, Erfan Yazdandoost Hamedani  
HIGHLIGHT: We consider cooperative multi-agent consensus optimization problems over an undirected network of agents, where only those agents connected by an edge can directly communicate.
- 565, TITLE: Fundamental Limits of Budget-Fidelity Trade-off in Label Crowdsourcing  
<https://papers.nips.cc/paper/6273-fundamental-limits-of-budget-fidelity-trade-off-in-label-crowdsourcing>  
AUTHORS: Farshad Lahouti, Babak Hassibi  
HIGHLIGHT: In this work, the CS problem, as a human-in-the-loop computation problem, is modeled and analyzed in an information theoretic rate-distortion framework.
- 566, TITLE: An Online Sequence-to-Sequence Model Using Partial Conditioning  
<https://papers.nips.cc/paper/6594-an-online-sequence-to-sequence-model-using-partial-conditioning>  
AUTHORS: Navdeep Jaitly, Quoc V. Le, Oriol Vinyals, Ilya Sutskever, David Sussillo, Samy Bengio  
HIGHLIGHT: In this paper, we present a Neural Transducer that can make incremental predictions as more input arrives, without redoing the entire computation.

567, TITLE: Learning to Poke by Poking: Experiential Learning of Intuitive Physics  
<https://papers.nips.cc/paper/6113-learning-to-poke-by-poking-experiential-learning-of-intuitive-physics>  
AUTHORS: Pulkit Agrawal, Ashvin V. Nair, Pieter Abbeel, Jitendra Malik, Sergey Levine  
HIGHLIGHT: We propose a novel approach based on deep neural networks for modeling the dynamics of robot's interactions directly from images, by jointly estimating forward and inverse models of dynamics.

568, TITLE: Learning Deep Parsimonious Representations  
<https://papers.nips.cc/paper/6263-learning-deep-parsimonious-representations>  
AUTHORS: Renjie Liao, Alex Schwing, Richard Zemel, Raquel Urtasun  
HIGHLIGHT: In this paper we aim at facilitating generalization for deep networks while supporting interpretability of the learned representations.

569, TITLE: Only H is left: Near-tight Episodic PAC RL  
<https://papers.nips.cc/paper/6052-only-h-is-left-near-tight-episodic-pac-rl>  
AUTHORS:  
HIGHLIGHT: In this paper, we present a PAC algorithm for reinforcement learning in episodic finite MDPs with time-dependent transitions that acts epsilon-optimal in all but  $O(S A H^3 / \epsilon \log(1 / \delta))$  episodes.