Interesting posters at NeurIPS'18

Yusuf Roohani

December 2018

Posters visited at NeurIPS 2018 held in Montreal, Canada. This selection reflects my interests and accounts for roughly 10% of all posters presented. Selected papers are from all 6 poster sessions. I created the section headings arbitrarily - several papers could fall under multiple headings.

1 Representation Learning

1.1 General

1. Disconnected Manifold Learning for Generative Adversarial Networks
2. Learning a latent manifold of odor representations from neural responses in piriform cortex
3. Point process latent variable models of larval zebrafish behavior
4. ConditionalGANS robust to noisy labels
5. BourGAN: Generative Networks with Metric Embeddings
6. MiME: Multilevel Medical Embedding of Electronic Health Records for Predictive Healthcare
7. Generative modeling for protein structures
8. Manifold Structured Prediction
9. Supervising Unsupervised Learning
10. Hamiltonian Variational Auto-Encoder
11. Delta-encoder: an effective sample synthesis method for few-shot object recognition spotlight talk worth reading for few shot learning ideas
12. Weakly Supervised Dense Event Captioning in Videos - text adaptive GANs - read
13. Text-Adaptive Generative Adversarial Networks: Manipulating Images with Natural Language

14. Variational PDEs for Acceleration on Manifolds and Application to Diffeomorphisms Variational PDE's for Acceleration on Manifolds
15. Learning Latent Subspaces in Variational Autoencoders - learning latent subspace in VAE
16. Representation Learning for Treatment Effect Estimation from Observational Data - representational learning for treatment effect estimation
17. https://nips.cc/Conferences/2018/Schedule?showEvent=11271 - simpleE embedding for link prediction in knowledge graphs
18. Watch Your Step: Learning Node Embeddings via Graph Attention learning node embeddings via Graph Attention read
19. Representation Learning of Compositional Data
20. Learning latent variable structured prediction models with Gaussian perturbations
21. Incorporating Context into Language Encoding Models for fMRI context in language encoding models for fMRI
22. Multi-value Rule Sets for Interpretable Classification with Feature-Efficient Representations- multi value rule sets for interpretable classification with feature efficient representations
23. Communication Efficient Parallel Algorithms for Optimization on Manifolds- Parallel algorithms for optimization on manifolds
24. Gaussian Process Prior Variational Autoencoders read
25. Learning semantic similarity in a continuous space - Learning semantic similarity in a continuous space

1.2 Graph representation learning

26. Navigating with Graph Representations for Fast and Scalable Decoding of Neural Language Models
27. Adaptive Sampling Towards Fast Graph Representation Learning
28. Hierarchical Graph Representation Learning with Differentiable Pooling
29. Beyond Grids: Learning Graph Representations for Visual Recognition *** read
30. Mean-field theory of graph neural networks in graph partitioning - read
31. Found Graph Data and Planted Vertex Covers

32. Graph Convolutional Policy Network for Goal-Directed Molecular Graph Generation

33. On Learning Markov Chains On learning markov chains read

34. Robust Learning of Fixed-Structure Bayesian Networks robust learning of fixed structure bayesian networks

35. DAGs with NO TEARS: Continuous Optimization for Structure Learning DAGs with NO Tears read

36. Symbolic Graph Reasoning Meets Convolutions - symbolic graph reasoning meets convolutions

37. Link Prediction Based on Graph Neural Networks - link prediction based on GNN

38. Constructing Deep Neural Networks by Bayesian Network Structure Learning - constructing DNN by Bayesian Structure Learning read

39. Constrained Graph Variational Autoencoders for Molecule Design - constrained graph VAE for Molecule Design read

40. Cluster Variational Approximations for Structure Learning of Continuous-Time Bayesian Networks from Incomplete Data- Structure learning of continuous time BNet - read

41. Combinatorial Optimization with Graph Convolutional Networks and Guided Tree Search- combinatorial optimization with gaps CN

42. Bayesian Structure Learning by Recursive Bootstrap - bayesian structure learning by recursive bootstrap

1.3 Domain adaptation, disentanglement

43. Unsupervised Image-to-Image Translation Using Domain-Specific Variational Information Bound

44. Unsupervised Attention-guided Image-to-Image Translation

45. Conditional Adversarial Domain Adaptation

46. Bayesian multi-domain learning for cancer subtype discovery from next-generation sequencing count data

47. Unsupervised Cross-Modal Alignment of Speech and Text Embedding Spaces

48. Extracting Relationships by Multi-Domain Matching multi domain matching - read sounds simple

49. A Unified Feature Disentangler for Multi-Domain Image Translation and Manipulation

50. Adaptive deep embeddings - A Synthesis of Methods for k-Shot Inductive Transfer Learning weight transfer + deep metric learning + few-shot learning. weight trasnfer is the least effective method for inductive transfer learning. histogram loss is robust regardless of the amount of label data in the target domaininterestingly similar to weakly supervised work

51. Insights on representational similarity in neural networks with canonical correlation

52. Generalizing to Unseen Domains via Adversarial Data Augmentation - read

53. Generalization Bounds for Uniformly Stable Algorithms

54. Revisiting (ϵ, γ, τ)-similarity learning for domain adaptation - Revisiting epsilon, gamma, tau similarity learning for domain adaptation

55. Algorithms and Theory for Multiple-Source Adaptation

56. Synthesize Policies for Transfer and Adaptation across Tasks and Environments - Policies for transfer and a adaptation across tasks and environments

57. Life-Long Disentangled Representation Learning with Cross-Domain Latent Homologies Disentangled representation learning with cross-domain latent homologies read

58. Isolating Sources of Disentanglement in Variational Autoencoders Isolating sources of entanglement in VAE

59. Image-to-image translation for cross-domain disentanglement - I2I translation for cross-domain disentanglement - readd

60. Transfer Learning with Neural AutoML - transfer learning with neuroautoml

61. Learning Deep Disentangled Embeddings With the F-Statistic Loss - learning deep disentangled embedding with f static loss

62. Supervised autoencoders: Improving generalization performance with unsupervised regularizers - supervised autoencoders

63. Domain-Invariant Projection Learning for Zero-Shot Recognition - Projection learning for zero shot recognition

64. A Simple Unified Framework for Detecting Out-of-Distribution Samples and Adversarial Attacks - Detecting out of distribution samples (know when you don't know)

2 Causal graphs

65. Experimental Design for Cost-Aware Learning of Causal Graphs - cost-aware learning of causal graphs - read
66. Domain Adaptation by Using Causal Inference to Predict Invariant Conditional Distributions- domain adaptation using causal inference read
67. Causal Inference with Noisy and Missing Covariates via Matrix Factorization- causal inference with noisy and missing covariates
68. Learning and Testing Causal Models with Interventions- Learning and testing causal models with interventions
69. Multi-domain Causal Structure Learning in Linear Systems- multi domain causal structure learning in linear systems - read
70. Direct Estimation of Differences in Causal Graphs- Direct estimation of difference in causal graphs - read
71. Identification and Estimation of Causal Effects from Dependent Data-read
72. Submodular Field Grammars: Representation, Inference, and Application to Image Parsing- Submodular field grammars

3 Practical reference

73. Size-Noise Tradeoffs in Generative Networks
74. Visualizing the Loss Landscape of Neural Nets
75. Bias and Generalization in Deep Generative Models: An Empirical Study If model see red car and blue bue, will it produce a red bus, built a framework to study these problems
76. Step size matters in deep learning
77. How Does Batch Normalization Help Optimization?
78. Sanity Checks for Saliency Maps
79. On Binary Classification in Extreme Regions
80. How to Start Training: The Effect of Initialization and Architecture
81. On GANs and GMMs On GANs and GMMs
82. Assessing Generative Models via Precision and Recall Assessing generative models via precision and recall read

83. When do random forests fail?
84. Learning to Multitask - learning to multitask
85. Towards Understanding Learning Representations: To What Extent Do Different Neural Networks Learn the Same Representation - Are intermediate representations robust - read
86. How to tell when a clustering is (approximately) correct using convex relaxations
87. Overfitting or perfect fitting? Risk bounds for classification and regression rules that interpolate- overfitting or perfect fitting read
88. Realistic Evaluation of Deep Semi-Supervised Learning Algorithms- realistic semi supervised learning read
89. Informative Features for Model Comparison - informative features for model comparison
90. Understanding Batch Normalization- understanding batch normalization - read
91. How Many Samples are Needed to Estimate a Convolutional Neural Network?- read results
92. Discrimination-aware Channel Pruning for Deep Neural Networks- NN pruning read bib
93. Implicit Bias of Gradient Descent on Linear Convolutional Networks

4 Theory, Optimization

94. Generalizing Point Embeddings using the Wasserstein Space of Elliptical Distributions
95. On Neuronal capacity Deeper architectures do not provide greater capacity but have better properties like regularization, smoother functions. New system for determining the capacity of an architecture. upper bound on the total number of bits that can be learnt form the data
96. Neural Ordinary Differential Equations ODE paper Traditional ODE solver takes small steps. Close connection between Resnet and Euler integrators. Taking the exact derivative through an ODE solution rather than backprop. It has O(1) memory gradients. Can now replace any set of resnet layers with ODE layers - with fewer paramaters than ResNet. Can train physics style models and time series models. Have a well defined state at all times. Biggest benefit turned out to be density modeling. FFJORD: Free-form Continuous Dynamics for Scalable Reversible Generative Models. New family of continuous depth architectures. Adaptive computation

97. Importance Weighting and Variational Inference

98. A theory on the absence of spurious solutions for nonconvex and nonsmooth optimization

99. The Limit Points of (Optimistic) Gradient Descent in Min-Max Optimization

100. Neural Arithmetic Logic Units - nalu is interesting

101. A Theory-Based Evaluation of Nearest Neighbor Models Put Into Practice- Theory based evaluation of Nearest neighbors models read

102. Sharp Bounds for Generalized Uniformity Testing - bounds for generalized uniformity testing

103. Scalable Robust Matrix Factorization with Nonconvex Loss - Scalable robust MF with non convex loss

104. Adversarially Robust Optimization with Gaussian Processes Adversarially robust optimization with gaussian processes

105. First-order Stochastic Algorithms for Escaping From Saddle Points in Almost Linear Time - First order stochastic algorithms for escaping saddle points in linear time

106. Connecting Optimization and Regularization Paths- connecting optimization and regularization paths read

107. The committee machine: Computational to statistical gaps in learning a two-layers neural network

108. Compact Representation of Uncertainty in Clustering

5 Miscellaneous

109. Leveraging the Exact Likelihood of Deep Latent Variable Models Enables quantitative comparison between architectures. Leads to the notion of structural regularization

110. ChannelNets: Compact and Efficient Convolutional Neural Networks via Channel-Wise Convolutions

111. Representer Point Selection for Explaining Deep Neural Networks

112. Multi-Task Learning as Multi-Objective Optimization

113. Revisiting Multi-Task Learning with ROCK: a Deep Residual Auxiliary Block for Visual Detection

114. Partially-Supervised Image Captioning
115. Neural Nearest Neighbors Networks
116. A Likelihood-Free Inference Framework for Population Genetic Data using Exchangeable Neural Networks
117. High Dimensional Linear Regression using Lattice Basis Reduction - hi dimensional linear regression read
118. l1-regression with Heavy-tailed Distributions
119. A Probabilistic U-Net for Segmentation of Ambiguous Images
120. On Coresets for Logistic Regression
121. Gather-Excite: Exploiting Feature Context in Convolutional Neural Networks Exploring feature context in CNNs read
122. Coupled Variational Bayes via Optimization Embedding Variational Bayes via Optimization Embedding
123. Sparse Covariance Modeling in High Dimensions with Gaussian Processessparse covariance modeling in high dimensions
124. Causal Discovery from Discrete Data using Hidden Compact Representation
125. Dynamic Network Model from Partial Observations dynamic network model from partial observations
126. Semi-supervised Deep Kernel Learning: Regression with Unlabeled Data by Minimizing Predictive Variance
127. Multitask Boosting for Survival Analysis with Competing Risks - multitask boosting for survival analysis
128. Binary Classification from Positive-Confidence Data - binary classification with confidence data
129. The Sparse Manifold Transform - Sparse manifold theorem
130. Manifold-tiling Localized Receptive Fields are Optimal in Similarity-preserving Neural Networks- localized receptive fields for similarity preserving rural networks
131. Integrated accounts of behavioral and neuroimaging data using flexible recurrent neural network models
132. Deep Homogeneous Mixture Models: Representation, Separation, and Approximation
133. Model-based targeted dimensionality reduction for neuronal population data Model based targeted dimensionality reduction for neuronal population data read

6 From talks

134. The tradeoffs of large scale learning Test of time award
135. Optimization methods for large-scale machine learning
136. Data science is science's second chance to get causal inference right: A classification of data science tasks
137. Clinically applicable deep learning for diagnosis and referral in retinal disease - Nature Medicine