Rui Shu

CONTACT INFORMATION

E-mail: ruishu@stanford.edu

Website: ruishu.io

Github: github.com/RuiShu

OVERVIEW

I am a CS Ph.D. student at Stanford University. My research focus has primarily been on domain adaptation, semi-supervised learning, and density estimation. I am interested in developing learning algorithms through the lens of regularization and seek to better understand the inductive bias of deep learning models.

EDUCATION

Stanford Unversity, Ph.D. (2017 – Present)
- Computer Science: Machine Learning/Artificial Intelligence
- Advisor: Stefano Ermon

Stanford University, M.Sc. (2015 – 2017)
- Biomedical Informatics. GPA: 3.92.
- Selected coursework: Representation learning in computation vision · probabilistic graphical models · sequential decision making · convex optimization

Dartmouth College, B.A. (2011 – 2014)
- Chemistry, with Minor in Statistics. GPA: 3.95
- Selected coursework: Theoretical machine learning · communication protocols and complexity · probability and statistical inference

WORK EXPERIENCE

DeepMind for Google, Mountain View, California USA
Research Intern, (2018)
In-progress. Currently developing new methods for domain adaptation and variational inference.

Adobe Research, San Jose, California USA
Research Intern, (2017 – 2018)
Developed a novel approach to unsupervised domain adaptation by leveraging the cluster assumption. Demonstrated that the model achieves state-of-the-art on several visual domain adaptation benchmarks. Work presented at the International Conference on Learning Representations (ICLR).

Adobe Research, San Jose, California USA
Research Intern, (2016 – 2017)
Developed a novel framework for semi-supervised high-dimensional conditional density estimation by hybridizing joint and conditional variational autoencoders. Work presented at the International Conference on Machine Learning (ICML).

Fliptop, San Francisco, California USA (acquired by LinkedIn)
Data Scientist Intern, (2015)
Created a Python pipeline for large-scale feature extraction and model evaluation. Demonstrated that simpler, well-regularized models were able to achieve 5% predictive accuracy improvement in comparison to Fliptop's existing algorithms.

PUBLICATIONS

R. Shu, H. Bui, S. Zhao, M. Kochenderfer, S. Ermon. Amortized Inference Regularization. In *Neural Information Processing Systems (NIPS)*, 2018.

Y. Song, **R. Shu**, Nate Kushman, S. Ermon. Generative Adversarial Examples. In *Neural*

Information Processing Systems (NIPS), 2018.

R. Shu, H. Bui, H. Narui, S. Ermon. A DIRT-T Approach to Unsupervised Domain Adaptation. In *International Conference on Learning Representations (ICLR)*, 2018.

S. Eismann, D. Levy, **R. Shu**, S. Ermon. Bayesian optimization and attribute adjustment. In *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2018.

E. Banijamali, **R. Shu**, M. Ghavamzadeh, H. Bui, and A. Ghodsi. Robust Locally-Linear Controllable Embedding. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017.

R. Shu, H. Bui, M. Ghavamzadeh. Bottleneck Conditional Density Estimation. In *International Conference on Machine Learning (ICML)*, 2017.

J. Brofos, **R. Shu**, and F. Zhang. The Optimistic Method for Model Estimation. In *International Symposium on Intelligent Data Analysis (IDA)*, 2016.

J. Brofos, **R. Shu**. Parallelization of Minimum Probability Flow on Binary Markov Random Fields. In *IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2015. (Best poster award)

P. Gurel, M. A, B. Guo, **R. Shu**, D. Mierke, H. Higgs. Assembly and Turnover of Short Actin Filaments by the Formin INF2 and Profilin. In *Journal of Biological Chemistry*, 2015.

B. Guo, P. Gurel, **R. Shu**, H. Higgs, M. Pellegrini, D. Mierke. Monitoring ATP hydrolysis and ATPase inhibitor screening using 1H NMR. In *Chemical Communications*, 2014.

P. Gurel, P. Ge, E. Grintsevich, **R. Shu**, L. Blanchoin, H. Zhou, E. Reisler, H. Higgs. INF2-Mediated Severing through Actin Filament Encirclement and Disruption. In *Cell*, 2014.

A. Shcheglovitov, O. Shcheglovitova, M. Yazawa, T. Portmann, **R. Shu**, V. Sebastiano, A. Krawisz, W. Froehlich, J. Bernstein, J. Hallmayer, R. Dolmetsch. SHANK3 and IGF1 restore synaptic deficits in neurons from 22q13 deletion syndrome patients. In *Nature*, 2013.

WORKSHOPS

R. Shu, H. Bui, S. Ermon. AC-GAN Learns a Biased Distribution. In *Neural Information Processing Systems (NIPS) Workshop on Bayesian Deep Learning*, 2017.

R. Shu, J. Brofos, F. Zhang, M. Ghavamzadeh, H. Bui, and M. Kochenderfer. Stochastic Video Prediction with Conditional Density Estimation. In *European Conference on Computer Vision (ECCV) Workshop on Action and Anticipation for Visual Learning*, 2016.

J. Brofos, **R. Shu**, M. Jinn, and M. Downs. Leveraging Deep Neural Networks as Kernels for Survival Analysis. In *Neural Information Processing Systems (NIPS) Workshop on Machine Learning in Healthcare*, 2015.

REVIEW CONTRIBUTIONS

Neural Information Processing Systems (NIPS). Top-218 reviewer.	(2018)
Advances on Approximate Bayesian Inference (AABI)	(2017)
Association for the Advancement of Artificial Intelligence (AAAI)	(2017)

Open-Source Projects

Available on github

Tensorbayes. A light-weight extension of TensorFlow designed to be an analog to Lasagne's Parmesan. Applied to several downstream generative model and deep learning projects.

ACGAN-Biased. Showed that the AC-GAN objective is a Lagrangian to an objective that constrains the generator from placing density near the decision boundary of the auxiliary classifier. Empirically verified that AC-GAN learns a biased distribution.

VAE-Clustering. Showed that M2 model can be reparameterized as a Gaussian Mixture Variational Autoencoder. Showed that, by constraining the generator and/or changing the variational inference procedure, the model can learn better clusters.

Fast-Style-Transfer. Yet another amortized style transfer implementation in TensorFlow. Improved upon Logan Engstrom's awesome repository by replacing transpose convolutions and adding TensorBoard visualizations.

Variational-Autoencoder Provided a design paradigm for the training of variational autoencoders in Torch. Applied framework to experiments on video prediction and multi-modal density estimation.

Automated-Statistician. Leveraged Gaussian Processes and reinforcement learning to build an automated system that performed predictive model and hyperparameter selection in a multiple-model setting.

Minimum-Probability-Flow-Learning. Extended Sohl-Dickstein's work on minimum probability flow by incorporating the use of graph factorization/auxiliary Markov random fields for parameter-estimation in binary pair-wise Markov random fields.

Neural-Net-Bayesian-Optimization. Implemented a distributed version of a Bayesian optimization framework that used a deep neural network as the surrogate model (based on Ryan Adams' work on scalable Bayesian optimization).

Masters and Undergraduate Research Experience

Stanford Intelligent Systems Laboratory
Independent researcher, (2016)
Applied conditional variational autoencoders to video prediction. Improved performance by using mixture of Gaussians as the latent variable distribution. Presented as a workshop contribution at ECCV Action and Anticipation for Visual Learning.

Stanford University
Research Assistant, (2015 - 2017)
Applied variational autoencoders and generative adversarial networks to semi-supervised anomaly detection of bone microfractures in radiology images. Demonstrated that semi-supervised deep generative models can be successfully applied to radiology image classification.

Dartmouth College
Research Assistant, (2013 - 2014)
Discovered unique response of the protein INF2 to the energy-storage molecule ATP. Demonstrated that the actin protein can either be assembled or disassembled using INF2 by changing the concentration of ATP. Work presented in Cell.

Stanford University, Stanford, California USA
Research Assistant, (2012)
Correlated gene expression with cellular and electrophysiological features in neurons from Phelan-McDermid syndrome patients. Work presented in Nature.

TEACHING EXPERIENCE	**Dartmouth College** **Teaching Science Fellow**, (2014 – 2015) Applied machine learning techniques to predict student performance in science classes based on prior academic indices. Provided resources for general chemistry education in the Dartmouth Chemistry Department. Managed a site devoted to correspondence between the Academic Skills Center, Dean's Office, and Teaching Science Fellows.
PROGRAMMING SKILLS	Deep learning software: PyTorch, TensorFlow, Keras, Torch, Theano Languages: Python, Java, C Scientific computing: R, MATLAB
HONORS AND AWARDS	John G. Kemeny Computing Prize—honorable mention (2015) Phi Beta Kappa Honor Society—associate member. Membership based on GPA (2014) Duke Data Science Competition—honorable mention (2014) Rofus Choate Scholar Award. Membership based on GPA (2012, 2013) Dartmouth Presidential Scholarship Award (2013)
GRANTS	Stanford Biomedical Informatics Travel Grant (2015) Neukom Institute Travel Grant (2014) Dartmouth Presidential Scholarship Award (2013) Dartmouth Undergraduate Leave Term Research (2013) Dartmouth Research Fellowship (2012)