Rui Shu

CONTACT
INFORMATION

E-mail: rui@openai.com

Website: ruishu.io
Github: github.com/RuiShu

OVERVIEW

I am a research scientist at OpenAI, specializing in the development of next-generation generative models with strong understanding and reasoning capabilities. Previously, I worked on variational methods and representation learning in the context of deep generative modeling, domain adaptation, and reinforcement learning.

EDUCATION

Ph.D. Computer Science, Stanford University (2017 − 2022)
Advisor: Stefano Ermon
M.Sc. Biomedical Informatics, Stanford University (2015 − 2017)
B.A. Chemistry, Minor in Statistics, Dartmouth College (2011 − 2014)

EMPLOYMENT

OpenAI, Research Scientist (2022 − Present)
Citadel Securities, Quantitative Research Intern (Jun − Aug 2020)
Google Brain, Research Intern (Jun − Sep 2019)
DeepMind for Google, Research Intern (Jun − Sep 2018)
Adobe Research, Research Intern (2017 − 2018)
Adobe Research, Research Intern (2016 − 2017)
Fliptop (acquired by LinkedIn), Data Scientist Intern (Jun − Sep 2015)

CONFERENCE
PUBLICATIONS

[1] **R. Shu**, S. Ermon. Bit Prioritization in Variational Autoencoders via Progressive Coding. In *International Conference on Machine Learning (ICML)*, 2022.

[2] T. Nguyen*, **R. Shu***, T. Pham, H. Bui, S. Ermon. Temporal Predictive Coding For Model-Based Planning In Latent Space. In *International Conference on Machine Learning (ICML)*, 2021.

[3] Y. Xu, Y. Song, S. Garg, L. Gong, **R. Shu**, A. Grover, S. Ermon. Anytime Sampling for Autoregressive Models via Ordered Autoencoding. In *International Conference on Learning Representations (ICLR)*, 2021.

[4] **R. Shu***, T. Nguyen*, Y. Chow, T. Pham, K. Than, M. Ghavamzadeh, S. Ermon, H. Bui. Predictive Coding for Locally-Linear Control. In *International Conference on Machine Learning (ICML)*, 2020.

[5] A. Grover, K. Choi, **R. Shu**, S. Ermon. Fair Generative Modeling via Weak Supervision. In *International Conference on Machine Learning (ICML)*, 2020.

[6] **R. Shu**, Y. Chen, A. Kumar, S. Ermon, B. Poole. Weakly Supervised Disentanglement with Guarantees. In *International Conference on Learning Representations (ICLR)*, 2020.

[7] N. Levine, Y. Chow, **R. Shu**, A. Li, M. Ghavamzadeh, H. Bui. Prediction, Consistency, Curvature: Representation Learning for Locally-Linear Control. In *International Conference on Learning Representations (ICLR)*, 2020.

[8] A. Grover, C. Chute, **R. Shu**, Z. Cao, S. Ermon. AlignFlow: Cycle Consistent Learning from Multiple Domains via Normalizing Flows. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2019.

[9] **R. Shu**, H. Bui, J. Whang, S. Ermon. Buffered Stochastic Variational Inference. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2019.

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

[11] Y. Song, **R. Shu**, Nate Kushman, S. Ermon. Generative Adversarial Examples. In *Neural Information Processing Systems (NeurIPS)*, 2018.

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

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

[14] 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.

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

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

[17] 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)

JOURNAL
PUBLICATIONS

[18] 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.

[19] 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.

[20] 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.

[21] 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.

WORKSHOP
PUBLICATIONS

[22] James Brofos, **Rui Shu**, Roy Lederman. A Bias-Variance Decomposition for Bayesian Deep Learning. In *Neural Information Processing Systems (NeurIPS) Workshop on Bayesian Deep Learning*, 2019.

[23] **R. Shu**, S. Zhao, M. Kochenderfer. Rethinking Style and Content Disentanglement in Variational Autoencoders. In *International Conference on Learning Representations (ICLR) Workshop*, 2018.

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

[25] **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.

Tensorsketch
A light-weight library for deep learning in TensorFlow 2.0.

Tensorbayes
A light-weight library for generative modeling and deep learning.

ACGAN-Biased
Empirically verified that AC-GAN learns a biased distribution.

VAE-Clustering
Clustering with Gaussian Mixture Variational Autoencoder.

Fast-Style-Transfer. Yet another amortized style transfer implementation in TensorFlow.

Variational-Autoencoder
Torch implementation for video prediction and density estimation.

Automated-Statistician
Gaussian Processes for automatic hyperparameter selection in a multiple-model setting.

Minimum-Probability-Flow-Learning
Extends minimum probability flow via auxiliary Markov random fields for parameter-estimation.

Neural-Net-Bayesian-Optimization
Distributed version of a Bayesian optimization framework that used a deep neural network.