

SHENGYANG SUN

University of Toronto ssy@cs.toronto.edu (+1) 647-879-4235

EDUCATION

University of Toronto (Department of Computer Science) PhD Candidate, Advised by Roger Grosse	Toronto, ON, Canada Sept. 2017 – present
Tsinghua University (Department of Electronic Engineering) Bachelor of Engineering	Beijing, China Sept. 2013 – Jul. 2017

RESEARCH INTERESTS

[Personal Page](#)

Machine Learning, Bayesian methods, Continual Learning

My research focuses on machine learning, especially the combination between Bayesian methods and deep neural networks. I aim to leverage probabilistic methods to improve the quality, reliability and efficiency of machine learning systems. Specifically, I investigate how to provide uncertainty estimation in probabilistic models and exploit the uncertainty to improve the robustness and guide exploration. Furthermore, I am interested in improving the learning efficiency and out-of-distribution generalization of intelligent systems, such as in continual learning and meta-learning. I am enthusiastic at both fundamental and applied research for solving realistic problems.

PUBLICATIONS

[Google Scholar](#)

Peer-reviewed Conference Publications

1. **Shengyang Sun**, Daniele Calandriello, Huiyi Hu, Ang Li, Michalis Titsias. Information-theoretic Online Memory Selection for Continual Learning. ICLR2022.
2. Jimmy Ba, Murat A Erdogdu, Marzyeh Ghassemi, **Shengyang Sun**, Taiji Suzuki, Denny Wu, Tianzong Zhang. Understanding the Variance Collapse of SVGD in High Dimensions. ICLR 2022.
3. **Shengyang Sun**, Jiaxin Shi, Andrew Gordon Wilson, Roger Grosse. Scalable Variational Gaussian Processes via Harmonic Kernel Decomposition. ICML 2021.
4. Chaoqi Wang*, **Shengyang Sun***, Roger Grosse. Beyond Marginal Uncertainty: How Accurately can Bayesian Regression Models Estimate Posterior Predictive Correlations? AISTATS 2021 [Oral Presentation].
5. Jun Yang*, **Shengyang Sun***, Daniel Roy. Fast-rate PAC-Bayes Generalization Bounds via Shifted Rademacher Processes. NeurIPS 2019.
6. **Shengyang Sun***, Guodong Zhang*, Jiaxin Shi*, Roger Grosse. Functional variational Bayesian neural networks. ICLR 2019.
7. James Lucas, **Shengyang Sun**, Richard Zemel, Roger Grosse. Aggregated Momentum: Stability Through Passive Damping. ICLR 2019.
8. **Shengyang Sun**, Guodong Zhang, Chaoqi Wang, Wenyuan Zeng, Jiaman Li, and Roger Grosse. Differentiable compositional kernel learning for Gaussian processes. ICML 2018.
9. Guodong Zhang*, **Shengyang Sun***, Roger Grosse. (2017). Natural Gradient as Stochastic Variational Inference. ICML 2018.
10. Jiaxin Shi, **Shengyang Sun**, Jun Zhu. (2017). A Spectral Approach to Gradient Estimation for Implicit Distributions. ICML 2018.

11. Jiaxin Shi*, **Shengyang Sun***, Jun Zhu. (2017). Kernel Implicit Variational Inference. ICLR 2018.
12. **Shengyang Sun**, Changyou Chan and Lawrence Carin. (2016). Learning Structured Weight Uncertainty in Bayesian Neural Networks. AISTATS 2017.

Preprints and Workshops

1. **Shengyang Sun***, Jiaxin Shi*, Roger Grosse. Neural Networks as Inter-domain Inducing Points. AABI 2021 Symposium.
2. Jiaxin Shi, Jianfei Chen, Jun Zhu, **Shengyang Sun**, Yucen Luo, Yihong Gu, Yuhao Zhou. ZhuSuan: A library for Bayesian deep learning. <https://arxiv.org/abs/1709.05870>.

Note: * represents equal contribution.

AWARDS

- Top Reviewers NeurIPS2020, ICLR2021
- Borealis AI Global Fellowship Award 2019
- Connaught New Researcher Award 2017
- Connaught International Scholarship (University-wide 20) 2017-2022
- First Class Undergraduate Scholarship 2014, 2015, 2016

EXPERIENCE

DeepMind London, UK
 Research Scientist Intern, Advisor: Michalis Titsias Jun. 2021 - Oct. 2021

- A challenging problem in task-free continual learning is the online selection of a representative replay memory from data streams. During the internship, I developed information-theoretic algorithms to tackle the online memory selection problem in task-free and data imbalanced continual learning.

Google AI Beijing, China
 Research Intern, Advisor: Chong Wang Jun. 2018 - Sept. 2018

- I studied how modern neural networks perform in terms of calibration, which terms how well the model's predictive probabilities are aligned to the underlying ground truth. I proposed approaches to increase the calibration of neural networks.

Duke University Durham, USA
 Research Scholar, Advisor: Lawrence Carin Jul. 2016 – Aug. 2016

- I proposed an efficient variational posterior that captures the correlations between individual weights for Bayesian neural networks. I demonstrated that the structured variational posterior lead to more accurate posterior approximations.

TALKS

- Information-theoretic Online Memory Selection @ Tsinghua University, 2021
- Harmonic variational Gaussian processes @ AI TIME, 2021
- Kernel Implicit Variational Inference @ Tsinghua University, 2018
- Functional variational Bayesian neural networks @ Google Toronto, 2018