

Pavel Izmailov

New York City, USA
[izmailovpavel.github.io](https://github.com/izmailovpavel)
izmailovpavel@gmail.com
[Google Scholar Profile](#)
[GitHub](#)

Education

- 2019–Now **PhD student, Courant Institute, New York University**
Computer Science department
- Supervisor: [Andrew Gordon Wilson](#)
 - Expected Graduation: May 2023
- 2017–2019 **PhD student, Cornell University**
Operations Research and Information Engineering department
- GPA: 3.9 / 4
 - Supervisor: [Andrew Gordon Wilson](#)
 - Obtained Masters degree and transferred to NYU
- 2013–2017 **BSc, Lomonosov Moscow State University**
Faculty of Computational Mathematics and Cybernetics
- GPA: 4.75 / 5

Experience

- 2021 Research Intern → Student Researcher, Google Research
- Supervisors: [Alex Alemi](#) and [Ben Poole](#)
 - Representation learning with information-constrained CLIP models (in progress)
- Summer 2020 Research Intern, Google Research
- Supervisor: [Matt Hoffman](#)
 - Large-scale Hamiltonian Monte Carlo for Bayesian Neural Networks
- Summer 2019 Research Intern, Amazon AWS
- Supervisors: [Yuyang \(Bernie\) Wang](#), [Alexander J. Smola](#)
 - Multi-scale time series forecasting
- 2015–2017 Undergraduate Student Researcher, [Bayesian Methods Research Group](#)
- Supervisors: [D. Vetrov](#), [D. Kropotov](#)
- 2016 Deep Learning Intern, [Artec 3D](#)
- Face recognition based on video data from 3D scanners

Publications

* Equal Contribution

- 2021 Dangers of Bayesian Model Averaging under Covariate Shift
ArXiv preprint
P. Izmailov, P. Nicholson, S. Lotfi, A. G. Wilson
[arxiv](#)
- 2021 Does Knowledge Distillation Really Work?
ArXiv preprint
S. Stanton, **P. Izmailov**, P. Kirichenko, A. A. Alemi, A. G. Wilson
[arxiv](#)
- 2021 What Are Bayesian Neural Network Posteriors Really Like?
International Conference on Machine Learning (ICML), **Long Talk (Oral)**
P. Izmailov, S. Vikram, M. D. Hoffman, A. G. Wilson
[arxiv](#)
- 2020 Learning Invariances in Neural Networks from Training Data
Neural Information Processing Systems (NeurIPS)
G. Benton, M. Finzi, **P. Izmailov**, A. G. Wilson
[arxiv](#)

- 2020 Why Normalizing Flows Fail to Detect Out-of-Distribution Data
Neural Information Processing Systems (NeurIPS)
P. Kirichenko*, **P. Izmailov***, A. G. Wilson
[arxiv](#)
- 2020 Bayesian Deep Learning and a Probabilistic Perspective of Generalization
Neural Information Processing Systems (NeurIPS)
A. G. Wilson, **P. Izmailov**
[arxiv](#)
- 2020 Generalizing Convolutional Neural Networks for Equivariance to Lie Groups on Arbitrary Continuous Data
International Conference on Machine Learning (ICML)
M. Finzi, S. Stanton, **P. Izmailov**, A. G. Wilson A. G. Wilson, **P. Izmailov**
[arxiv](#)
- 2020 Semi-Supervised Learning with Normalizing Flows
International Conference on Machine Learning (ICML)
P. Izmailov*, P. Kirichenko*, M. Finzi*, A. G. Wilson
[arxiv](#)
- 2019 A Simple Baseline for Bayesian Uncertainty in Deep Learning
Neural Information Processing Systems (NeurIPS)
W. Maddox*, T. Garipov*, **P. Izmailov***, D. Vetrov, A. G. Wilson
[arxiv](#)
- 2019 Subspace Inference for Bayesian Deep Learning
Uncertainty in Artificial Intelligence (UAI)
P. Izmailov*, W. Maddox*, P. Kirichenko*, T. Garipov*, D. Vetrov, A. G. Wilson
[arxiv](#)
- 2019 There Are Many Consistent Explanations of Unlabeled Data: Why You Should Average
International Conference on Learning Representations (ICLR)
B. Athiwaratkun, M. Finzi, **P. Izmailov**, A. G. Wilson
[OpenReview](#)
- 2018 Loss Surfaces, Mode Connectivity, and Fast Ensembling of DNNs
Neural Information Processing Systems (NeurIPS), **Spotlight Presentation**
T. Garipov*, **P. Izmailov***, D. Podoprikin*, D. Vetrov, A. G. Wilson
[arxiv](#)
- 2018 Averaging Weights Leads to Wider Optima and Better Generalization
Uncertainty in Artificial Intelligence (UAI), **Oral Presentation**
P. Izmailov*, D. Podoprikin*, T. Garipov*, D. Vetrov, A. G. Wilson
[arxiv](#)
- 2018 Scalable Gaussian Processes with Billions of Inducing Inputs via Tensor Train Decomposition
Artificial Intelligence and Statistics (AISTATS), **Oral Presentation**
P. Izmailov, A. Novikov, D. Kropotov
[arxiv](#)
- 2018 Tensor Train decomposition on TensorFlow (T3F)
Journal of Machine Learning Research (JMLR, published in 2020)
A. Novikov, **P. Izmailov**, V. Khrulkov, M. Fournov, I. Oseledets
[arxiv](#)
- 2017 Faster variational inducing input Gaussian process classification
Journal of Machine Learning and Data Analysis (JMLDA)
P. Izmailov, D. Kropotov
[arxiv](#)

Workshop Papers

- 2019 Subspace Inference for Bayesian Deep Learning
ICML Workshop on Uncertainty and Robustness in Deep Learning, **Oral Presentation**
P. Izmailov*, W. Maddox*, T. Garipov*, P. Kirichenko*, A. G. Wilson
[PDF](#)
- 2019 Semi-Supervised Learning with Normalizing Flows
ICML Workshop on Invertible Neural Nets and Normalizing Flows
P. Izmailov*, P. Kirichenko*, M. Finzi*, A. G. Wilson
[PDF](#)

- 2019 Invertible Convolutional Networks
*ICML Workshop on Invertible Neural Nets and Normalizing Flows, **Spotlight Presentation***
 M. Finzi*, **P. Izmailov***, W. Maddox*, P. Kirichenko*, A. G. Wilson
[PDF](#)
- 2018 Fast Uncertainty Estimates and Bayesian Model Averaging of DNNs
*UAI Workshop on Uncertainty in Deep Learning, **Oral Presentation***
 W. Maddox, T. Garipov, **P. Izmailov**, A. G. Wilson
[PDF](#)
- 2018 Improving Stability in Deep Reinforcement Learning with Weight Averaging
UAI Workshop on Uncertainty in Deep Learning
 E. Nikishin, **P. Izmailov**, B. Athiwaratkun, P. Shvechikov, D. Podoprikin, T. Garipov,
 D. Vetrov, A. G. Wilson
[PDF](#)

Talks

- 2021 Advances in Approximate Bayesian Inference, AABI (**Invited Talk**)
What Are Bayesian Neural Network Posteriors Really Like?
[Video](#)
- 2021 Max Plank Institute MIS and UCLA joint Seminar: Math Machine Learning
What Are Bayesian Neural Network Posteriors Really Like?
[Video](#)
- 2021 Teams at Google Brain and Perception
What Are Bayesian Neural Network Posteriors Really Like?
- 2021 Oxford Applied and Theoretical Machine Learning Group
What Are Bayesian Neural Network Posteriors Really Like?
- 2021 Teams at Google Brain, Translate and Perception
Does Knowledge Distillation Really Work?
- 2021 Oxford Applied and Theoretical Machine Learning Group
What Are Bayesian Neural Network Posteriors Really Like?
- 2021 International Conference on Machine Learning (ICML)
What Are Bayesian Neural Network Posteriors Really Like?
- 2019 Broad Institute of MIT and Harvard
How do we build neural networks we can trust?
[Video](#)
- 2018 Uncertainty in Artificial Intelligence (UAI)
Averaging Weights Leads to Wider Optima and Better Generalization
[Video](#)
- 2018 Artificial Intelligence and Statistics (AISTATS)
Scalable Gaussian Processes with Billions of Inducing Inputs via Tensor Train Decomposition

Awards

- 2021 Harold Grad Memorial Prize
 NYU Courant Institute prize for outstanding performance and promise as a graduate student
- 2019 NYU MacCracken PhD Fellowship
- 2019 NeurIPS highest-scoring reviewer award
- 2017-2018 Cornell University PhD Fellowship
- 2018 NeurIPS travel award and highest-scoring reviewer award
- 2018 AISTATS travel award
- 2018 UAI travel award
- 2017 Best undergraduate thesis award, Moscow State University CS department

Summer Schools

- 2017 [Summer school on NLP](#) by Moscow Institute of Physics and Technology
- 2017 [Spring school "Structural Inference"](#) by association of German universities
- 2017 [Winter school on Reinforcement Learning](#) by Moscow Institute of Physics and Technology

2014 [Rome-Moscow school of Matrix Methods and Applied Linear Algebra](#) by Moscow State University and University of Rome “Tor Vergata”

Teaching

2019 New York University
Teaching Assistant for “Bayesian Machine Learning” course

2018 Cornell University
Teaching Assistant for “Bayesian Machine Learning” course

Organizing

2021 Lead student organizer of the NeurIPS competition [“Approximate Inference in Bayesian Deep Learning”](#)

Reviewing

Conferences TMLR, ICML 2022, NeurIPS 2021, NeurIPS 2020, UAI 2020, ICML 2020, ICLR 2020, NeurIPS 2019 (Top 400 highest-scoring reviewers), UAI 2019, ICML 2019, UAI 2018, NeurIPS 2018 (Top 218 highest-scoring reviewers), AISTATS 2019, ICML 2019

Workshops ICML 2021 INNF workshop, ICML 2021 UDL workshop, NeurIPS 2019 Bayesian Deep Learning Workshop, ICML 2018 Workshop on Theoretical Foundations and Applications of Deep Generative Models

Skills

Proficient Python, SciPy stack, PyTorch, JAX, Git, L^AT_EX
Used before MXNet, Gluon, Tensorflow