

Polina Kirichenko

pol.kirichenko@gmail.com ✉
polkirichenko.github.io 🌐
Google Scholar, Semantic Scholar 🎓

Education

Ph.D. student in Data Science, New York University New York, USA
Center for Data Science; supervisor: Professor [Andrew Gordon Wilson](#) 2019 – current
Research interests: out-of-distribution generalization, robustness, uncertainty estimation

Ph.D. student in Operations Research, Cornell University Ithaca, USA
Operations Research and Information Engineering department; transferred to NYU 2018 – 2019

B.Sc. in Computer Science, Higher School of Economics Moscow, Russia
Computer Science department; supervisor: Professor [Dmitry Vetrov](#) 2014 – 2018
Cumulative GPA: 9.1 (10.0 scale), class rank: top 3%

Work Experience

Meta AI New York, USA
Visiting Scholar (FAIR-NYU AI Mentorship program); supervisor: [Rama Vedantam](#) Oct 2022 – Oct 2023
Research topic: out-of-distribution generalization

Meta AI Menlo Park, USA
Research Intern at AI Integrity team; supervisors: [Hamed Firooz](#), [Randall Balestrierio](#), [David Lopez-Paz](#) June 2022 – Sep 2022
Research topic: biases of data augmentation and regularization in deep networks

Cold Spring Harbor Laboratory Cold Spring Harbor, USA
Research Intern; supervisor: Prof. [Anthony Zador](#) June 2021 – Aug 2021
Research topics: meta-learning with compressed neural networks

DeepMind (remotely) Mountain View, USA
Research Scientist Intern; supervisors: [Mehrdad Farajtabar](#), [Razvan Pascanu](#), [Balaji Lakshminarayanan](#) June 2020 – Oct 2020
Research topic: continual learning with deep generative models

École Polytechnique Fédérale de Lausanne (EPFL) mlo.epfl.ch, Lausanne, Switzerland
Machine Learning and Optimization Lab June 2018 – Aug 2018
Research Intern; supervisors: Prof. [Martin Jaggi](#), Prof. [Dan Alistarh](#)
Research topic: evolution strategies for low precision training of neural networks

Bayesian Methods Research Group bayesgroup.ru, Moscow, Russia
Research Assistant; supervisor: Prof. Dmitry Vetrov Sep 2016 – Aug 2018
Research topic: structured sparsification of Bayesian neural networks

Google Seattle, USA
Software Engineering Intern, Google Cloud Platform Team July 2017 – Sep 2017

Google Munich, Germany
STEP Software Engineering Intern, Piper Team (Google's version control system) July 2016 – Sep 2016

Publications

* Equal Contribution

On Feature Learning in the Presence of Spurious Correlations [\[workshop pdf\]](#)
Pavel Izmailov*, **Polina Kirichenko***, Nate Gruver*, Andrew Gordon Wilson
First presented at *ICML Workshop on Principles of Distribution Shift 2022*
Neural Information Processing Systems (NeurIPS) 2022

Chroma-VAE: Mitigating Shortcut Learning with Generative Classifiers

Wanqian Yang, Polina Kirichenko, Micah Goldblum, Andrew Gordon Wilson

Neural Information Processing Systems (NeurIPS) 2022

Last Layer Re-Training is Sufficient for Robustness to Spurious Correlations [\[arXiv, code\]](#)

Polina Kirichenko*, Pavel Izmailov*, Andrew Gordon Wilson

ICML Workshop on Spurious Correlations, Invariance, and Stability 2022 (oral presentation)

Under Review

Does Knowledge Distillation Really Work? [\[arXiv, code\]](#)

Samuel Stanton, Pavel Izmailov, Polina Kirichenko, Alexander A. Alemi, Andrew G. Wilson

Neural Information Processing Systems (NeurIPS) 2021

Task-agnostic Continual Learning with Hybrid Probabilistic Models [\[arXiv, poster\]](#)

Polina Kirichenko, Mehrdad Farajtabar, Dushyant Rao, Balaji Lakshminarayanan, Nir Levine, Ang Li, Huiyi Hu, Andrew Gordon Wilson, Razvan Pascanu

ICML Workshop on Invertible Neural Networks and Normalizing Flows 2021 (spotlight talk)

Why Normalizing Flows Fail to Detect Out-of-Distribution Data [\[arXiv, code\]](#)

Polina Kirichenko*, Pavel Izmailov*, Andrew G. Wilson

First presented at *Workshop on Invertible Neural Networks and Normalizing Flows at ICML 2020*

Neural Information Processing Systems (NeurIPS) 2020

Semi-Supervised Learning with Normalizing Flows [\[arXiv, poster, code\]](#)

Pavel Izmailov*, Polina Kirichenko*, Marc Finzi*, Andrew G. Wilson

First presented at *Workshop on Invertible Neural Networks and Normalizing Flows at ICML 2019*

International Conference on Machine Learning (ICML) 2020

Subspace Inference for Bayesian Deep Learning [\[arXiv, poster, slides, code\]](#)

Pavel Izmailov*, Wesley Maddox*, Polina Kirichenko*, Timur Garipov*, Dmitry Vetrov, Andrew G. Wilson

First presented at *Workshop on Uncertainty & Robustness in Deep Learning at ICML 2019 (contributed talk)*

Uncertainty in Artificial Intelligence (UAI) 2019

SWALP: Stochastic Weight Averaging in Low Precision Training [\[PMLR, code\]](#)

Guandao Yang, Tianyi Zhang, Polina Kirichenko, Junwen Bai, Andrew G. Wilson, Christopher De Sa

International Conference on Machine Learning (ICML) 2019

Workshop Papers

Effective Surrogate Models for Protein Design with Bayesian Optimization [\[pdf\]](#)

Nate Gruver, Samuel Stanton, Polina Kirichenko, Marc Finzi, Phillip Maffettone, Vivek Myers, Emily Delaney, Peyton Greenside, Andrew Gordon Wilson

ICML Workshop on Computational Biology 2021

Invertible Convolutional Networks [\[pdf, poster\]](#)

Marc Finzi*, Pavel Izmailov*, Wesley Maddox*, Polina Kirichenko*, Andrew G. Wilson

Workshop on Invertible Neural Nets and Normalizing Flows at ICML 2019 (spotlight talk)

Awards

DeepMind Fellowship 2019

New York University Center for Data Science Graduate Fellowship 2019

Golden HSE Award (Alumni Success category) [\[link\]](#), 2019

HSE Alumni Academic Fellowship [\[link\]](#), 2019

NeurIPS Travel Award 2019

ICML Travel Award 2019

Cornell Operations Research and Information Engineering Graduate Fellowship 2018

Travel Grant for Women in Data Science Conference [\[link\]](#), 2018, 2019

Ilya Segalovich Scholarship (Yandex) [\[link\]](#), 2016, 2017

Google Anita Borg Memorial Scholarship (Women Techmakers Scholarship) 2015

Google Travel Grant for the Grace Hopper Celebration of Women in Computing 2015

Reviewing

Conferences: NeurIPS 2019 (top 400 highest-scoring reviewers), ICLR 2020, ICML 2020 (top 33% reviewer), UAI 2020, NeurIPS 2020, AISTATS 2021, AISTATS 2022, NeurIPS 2022

Workshops: NeurIPS 2019 WiML, NeurIPS 2019 BDL, ICML 2020 UDL, NeurIPS 2020 HAMLETS, ICML 2021 INN+ , ICML 2021 UDL, NeurIPS 2021 BDL

Talks

“Last Layer Re-Training is Sufficient for Robustness to Spurious Correlations”

- Oral Presentation at the ICML Workshop on Spurious Correlations, Invariance, and Stability [\[link\]](#), 2022
- Genentech, AI seminar 2022

“Robustness of Deep Learning Models to Distribution Shift”, WiML Un-Workshop at ICML 2022

“Why Normalizing Flows Fail to Detect Out-of-Distribution Data”

- ML Collective, Deep Learning: Classics and Trends (with Robin T. Schirrmeyer) [\[slides\]](#), 2021
- Facebook AI Research, Uncertainty team 2021
- CogSys Talks, Technical University of Denmark [\[video\]](#), 2020
- Capital One, Machine Learning seminar 2020
- NeurIPS 2020 [\[video\]](#), 2020
- INN+ : Invertible Neural Networks and Normalizing Flows workshop at ICML [\[video\]](#), 2020

“Applications of normalizing flows: semi-supervised learning, anomaly detection, and continual learning”

- **Keynote talk at ICML Workshop on Representation Learning for Finance Applications** [\[video\]](#), 2021

“Does your model know what it doesn’t know?”, WiML Un-Workshop at ICML 2021

“Task-agnostic Continual Learning with Hybrid Probabilistic Models”, ICML INN+ 2021 [\[video\]](#), 2021

“Continual Learning in Neural Networks” [in Russian], Bayesian Methods Research Group seminar [\[video\]](#), 2021

“Anomaly detection via Generative Models”

- ODS DafaFest 2020, Uncertainty Estimation in ML Workshop [\[video\]](#), 2020

“Uncertainty Estimation in Bayesian Deep Learning”, WiML Un-Workshop at ICML 2020

“Subspace Inference for Bayesian Deep Learning”

- University of Paris-Saclay, UQSay seminar [\[link\]](#), 2021
- Uncertainty and Robustness in Deep Learning workshop at ICML [\[video\]](#), 2019
- Higher School of Economics (with Pavel Izmailov) [\[video\]](#), 2019

“**How do we build neural networks we can trust?**”, **Broad Institute of MIT and Harvard** [\[video\]](#), 2019

Technical Skills

Programming languages: Python, C++

Deep learning frameworks: PyTorch, TensorFlow, Keras

Teaching

Cornell University Ithaca, USA

Teaching Assistant for “Data Science for All” course Jan 2019 – May 2019

Bayesian Methods for Machine Learning on Coursera [\[course link\]](#), Moscow, Russia

Teaching Assistant; helped prepare assignments and quizzes Sep 2017 – Aug 2018

The specialization of the course received Coursera [Outstanding Educator Award](#)

National Research University Higher School of Economics Moscow, Russia

Teaching Assistant for “Probability Theory and Statistics” (Sep 2016 – June 2017),

“Introduction to Data Analysis” (Jan 2016 – June 2016), “Introduction to Programming” (Sep 2015 – Dec 2015)