

# Ekaterina Lobacheva

[Home page](#) / [Google Scholar](#) / [GitHub](#)

e-mail: lobacheva.tjulja@gmail.com tel: +79166498371

I'm a deep learning researcher mainly focusing on understanding the properties of neural network training, loss landscape, and generalization. I am also interested in ensemble methods, properties of training dynamics and object representations obtained in different training paradigms, such as self-supervised and transfer learning, and scaling laws for neural networks. Currently, I am an independent researcher collaborating with [Bayesian Methods Research Group](#), where I did my PhD advised by [Dmitry Vetrov](#), and with [Mila](#), working with [Nicolas Le Roux](#) and [Irina Rish](#). The results of my work were published at NeurIPS, AAAI, EMNLP, ICCV.

## PUBLICATIONS

\* denotes joint first co-authorship

**To Stay or Not to Stay in the Pre-train Basin: Insights on Ensembling in Transfer Learning,** [paper](#)  
NeurIPS 2023

Ildus Sadrtidinov\*, Dmitrii Pozdeev\*, Dmitry Vetrov, **Ekaterina Lobacheva**

**Training Scale-Invariant Neural Networks on the Sphere Can Happen in Three Regimes,** [paper](#) / [code](#)  
NeurIPS 2022

Maxim Kodryan\*, **Ekaterina Lobacheva\***, Maksim Nakhodnov\*, Dmitry Vetrov

**On the Periodic Behavior of Neural Network Training with Batch Normalization and Weight Decay,** [paper](#) / [code](#)  
NeurIPS 2021

**Ekaterina Lobacheva\***, Maxim Kodryan\*, Nadezhda Chirkova, Andrey Malinin, Dmitry Vetrov

**On Power Laws in Deep Ensembles,** NeurIPS 2020 (Spotlight) [paper](#) / [code](#)  
**Ekaterina Lobacheva,** Nadezhda Chirkova, Maxim Kodryan, Dmitry Vetrov

**Structured Sparsification of Gated Recurrent Neural Networks,** AAAI 2020 (Oral) [paper](#) / [code](#)  
**Ekaterina Lobacheva\***, Nadezhda Chirkova\*, Alexander Markovich, Dmitry Vetrov

**Bayesian Compression for Natural Language Processing,** EMNLP 2018 [paper](#) / [code](#)  
Nadezhda Chirkova\*, **Ekaterina Lobacheva\***, Dmitry Vetrov

**Deep Part-Based Generative Shape Model with Latent Variables,** BMVC 2016 [paper](#)  
Alexander Kirillov, Mikhail Gavrikov, **Ekaterina Lobacheva,** Anton Osokin, Dmitry Vetrov

**Joint Optimization of Segmentation and Color Clustering,** ICCV 2015 [paper](#)  
**Ekaterina Lobacheva,** Olga Veksler, Yuri Boykov

## SELECTED WORKSHOP PUBLICATIONS AND PREPRINTS

**Large Learning Rates Improve Generalization: But How Large Are We Talking About?** [paper](#)  
NeurIPS 2023 M3L Workshop

**Ekaterina Lobacheva\***, Eduard Pockonechnyy\*, Maxim Kodryan, Dmitry Vetrov

**On the Memorization Properties of Contrastive Learning,** ICML 2021 OPPO Workshop [paper](#)  
Ildus Sadrtidinov, Nadezhda Chirkova, **Ekaterina Lobacheva**

**Deep Ensembles on a Fixed Memory Budget: One Wide Network or Several Thinner Ones?,** 2020 [paper](#)  
Nadezhda Chirkova, **Ekaterina Lobacheva,** Dmitry Vetrov

**Monotonic models for real-time dynamic malware detection,** ICLR Workshop 2018 [paper](#)  
Alexander Chistyakov, **Ekaterina Lobacheva,** Alexander Shevelev, Alexey Romanenko

## PROFESSIONAL EXPERIENCE

- 2023 - now      **Independent Researcher**  
I am currently working on an analysis of neural network training trajectories in the function space and related generalization metrics with [Nicolas Le Roux](#). Additionally, I am leading a project on ensembling in transfer learning setup and participating in a project on training with large learning rates at [Bayesian Methods Research Group](#).
- 2020 - 2022      **Research Fellow and Deputy Head, Centre of Deep Learning and Bayesian Methods, HSE University**  
I worked mostly on understanding the properties of neural networks training and loss landscape. Specifically, my work was focused on sharp and flat optima, mode connectivity, ensembling methods and specifics of training of networks with normalization layers. I also worked on the scaling laws for deep ensembles. The results were published at NeurIPS.
- 2018 - 2020      **Research Fellow, Samsung-HSE Laboratory at HSE University**  
I worked on Bayesian sparsification methods for recurrent neural networks, including embedding layers and gated layers such as LSTM. The results were published at EMNLP and AAAI.
- 2015 - 2018      **Junior Researcher, Kaspersky Lab**  
I worked on feature extraction techniques and classification models for dynamic malware detection. The results were published at ICLR Workshops and were built into company antivirus products.
- Summer 2014      **Research Intern, University of Western Ontario**  
I worked on the improvement of energy-based segmentation methods for the case of camouflage images. The results were published at ICCV. Advisors: [Yuri Boykov](#) and [Olga Veksler](#)

## EDUCATION

- 2022              **PhD in Computer Science, HSE University**  
**Thesis on:** Deep learning architectures on a fixed memory budget    **Advisor:** [Dmitry Vetrov](#)  
Cum Laude
- 2009 - 2014      **Specialist degree (BSc+MSc) in Computer Science, Lomonosov Moscow State University**  
**Thesis on:** Boltzmann machines for image segmentation    **Advisor:** [Dmitry Vetrov](#)  
Graduated with honors (GPA 5.0 out of 5.0)

## TECHNICAL SKILLS

- I mostly program in **Python** but also have some experience with C++
- I am fluent with common data science tools such as **NumPy**, **matplotlib**, **scikit-learn**, **pandas**
- My primary deep learning framework is **PyTorch** (prior to that, I had experience with Theano+Lasagne and TensorFlow)
- I'm comfortable with the common data science environment e.g., **bash**, **git**, **Linux**, **GPU clusters**

## THESIS SUPERVISION AND CO-SUPERVISION

- [Ildus Sadrtidinov](#) (now a PhD student at HSE University)
  - On the Memorization Properties of Contrastive Learning (BSc, 2021)
  - Ensembling Neural Networks in the Transfer Learning Setup (MSc, 2023)
- [Sergey Troshin](#) (now a PhD student at University of Amsterdam)
  - Deep Equilibrium ResNet (BSc, 2020)
- [Maksim Ryabinin](#) (PhD, now Distinguished Research Scientist at Together AI)
  - Gradient Optimization of Beam Search Hyperparameters (BSc, 2019)

- [Polina Kirichenko](#) (now a PhD student at NYU)
  - Study of Bayesian Regularization of Neural Networks (BSc, 2018)
- [Nadezhda Chirkova](#) (PhD, now Research Scientist at Naver Labs)
  - Bayesian Compression for Natural Language Processing, (MSc, 2018)

## TEACHING

2016 - Now	<b>Research seminar Machine Learning and Applications</b> , Faculty of Computer Science at HSE University
2015 - 2021	<b>Bayesian Methods in Machine Learning (organization + seminars)</b> , HSE University, MSU, and Yandex School of Data Analysis
2018 - 2019	<b>Neurobayesian models (organization + seminars)</b> , HSE University, MSU, and Yandex School of Data Analysis
2017	<b>Introduction to Deep Learning (Online course at Coursera, lectures)</b>
2016 - 2017	<b>Deep learning (lectures and seminars)</b> , MSU
2016 - 2017	<b>Machine learning (lectures)</b> , HSE and HSE/NES Programmes in Economics
2015 - 2017	<b>Data analysis (seminars)</b> , HSE University

Additionally, I was one of the main organizers of [Deep|Bayes Summer School](#) at 2017-2019 and gave several lectures and seminars there. Also, we with Nadezhda Chirkova gave a tutorial on Bayesian machine learning at [Machine Learning in High Energy Physics Summer School](#) (Hamburg, Germany, July 2019 and online, July 2020 and 2021) and [Workshop on Machine Learning and Applications to Physics](#) (Madrid, Spain, Dec 2019).

## PROGRAM COMMITTEE

Neural Information Processing Systems, NeurIPS (reviewer, 2019–2021, 2023):

- 2019: top-50% highest-scored reviewer
- 2021: outstanding reviewer award (top-8%)
- 2023: top reviewer