Ekaterina Lobacheva

Home page / Google Scholar / GitHub / Twitter
e-mail: lobacheva.tjulja@gmail.com tel: +1 438-365-3056

I'm a deep learning researcher mainly focusing on understanding the properties of neural network training and how they affect learned data representations and model generalization. I am also interested in ensemble and model-averaging methods, properties of neural network loss landscape, and specifics of training dynamics and representation learning in various training paradigms, such as self-supervised, transfer, continuous and online learning. The results of my work were published at top-tier conferences, including NeurIPS, AAAI, EMNLP, ICCV. Currently, I am a postdoc at Mila and Université de Montréal, working with Nicolas Le Roux and Irina Rish.

SELECTED PUBLICATIONS

[*] equal contribution, [†] shared senior authorship

Where Do Large Learning Rates Lead Us?, NeurIPS 2024 — paper
Ildus Sadrtdinov[*], Maxim Kodryan[*], Eduard Pokonechny[*], **Ekaterina Lobacheva**[†], Dmitry Vetrov[†]

To Stay or Not to Stay in the Pre-train Basin: Insights on Ensembling in Transfer Learning, NeurIPS 2023 — paper / code
Ildus Sadrtdinov[*], Dmitrii Pozdeev[*], Dmitry Vetrov, **Ekaterina Lobacheva**

Training Scale-Invariant Neural Networks on the Sphere Can Happen in Three Regimes, NeurIPS 2022 — paper / code
Maxim Kodryan[*], **Ekaterina Lobacheva**[*], Maksim Nakhodnov[*], Dmitry Vetrov

On the Periodic Behavior of Neural Network Training with Batch Normalization and Weight Decay, NeurIPS 2021 — paper / code
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

SELECTED WORKSHOP PUBLICATIONS AND PREPRINTS

How Learning Rates Shape Neural Network Focus: Insights from Example Ranking — paper
NeurIPS 2024 SciForDL Workshop
Ekaterina Lobacheva, Keller Jordan, Aristide Baratin, Nicolas Le Roux

Language Model Scaling Laws and Zero-Sum Learning, NeurIPS 2024 SciForDL Workshop — paper
Andrei Mircea, Nima Chitsazan, Supriyo Chakraborty, Renkun Ni, Austin Zhang,
Ekaterina Lobacheva, Irina Rish

Gradient Dissent in Language Model Training and Saturation, ICML 2024 HiLD Workshop — paper
Andrei Mircea, **Ekaterina Lobacheva**, Irina Rish

On the Memorization Properties of Contrastive Learning, ICML 2021 OPPO Workshop — paper
Ildus Sadrtdinov, Nadezhda Chirkova, **Ekaterina Lobacheva**

Monotonic models for real-time dynamic malware detection, ICLR Workshop 2018 — paper
Alexander Chistyakov, **Ekaterina Lobacheva**, Alexander Shevelev, Alexey Romanenko

PROFESSIONAL EXPERIENCE

2024 - *now* **Postdoctoral Research Fellow, Mila Quebec AI Institute and Université de Montréal**
My current projects are focused on the effects of training dynamics and hyperparameters on the properties of final models. More specifically, I investigate how the learning rate affects the network focus over the data distribution, how gradient opposition impacts foundational models at different scales, and how loss landscape properties change due to data distribution shifts in the transfer or continual learning settings. **Advisors:** Nicolas Le Roux and Irina Rish.

2023 - 2024 **Independent Researcher**
I led projects on ensembling in transfer learning setup and effects of large initial learning rate on network training, collaborating closely with researchers from Bayesian Methods Research Group. I also investigated neural network training trajectories in the function space and related generalization metrics in collaboration with Nicolas Le Roux.

2020 - 2022 **Research Fellow and Deputy Head, Centre of Deep Learning and Bayesian Methods, HSE University**
My research primarily focused on understanding the properties of neural networks training and loss landscape. Specifically, I investigated sharp and flat optima, mode connectivity, ensembling methods and the effects of normalization layers on training dynamics. The results were published at NeurIPS. As Deputy Head, I contributed to the organization and management of the lab.

2018 - 2020 **Research Fellow, Samsung-HSE Laboratory at HSE University**
I worked on Bayesian sparsification methods for recurrent neural networks, including techniques for embedding layers and gated layers such as LSTMs. The findings were published at EMNLP and AAAI.

2015 - 2018 **Junior Researcher, Kaspersky Lab**
I focused on feature extraction techniques and classification models for dynamic malware detection. The results were published at ICLR Workshops and built into the company's antivirus products.

Summer 2014 **Research Intern, University of Western Ontario**
I contributed to improving energy-based segmentation methods for 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** and my primary deep learning framework is **PyTorch**
- I am proficient with common data science tools such as **NumPy**, **matplotlib**, **scikit-learn**, **pandas**
- I'm comfortable with the common data science environment, including **bash**, **git**, **Linux**, **GPU clusters**

SELECTED THESIS SUPERVISION AND CO-SUPERVISION

- Ildus Sadrtdinov (currently 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 (currently a PhD student at University of Amsterdam)
  - Deep Equilibrium ResNet (BSc, 2020)

- Maksim Ryabinin (PhD, currently a Distinguished Research Scientist at Together AI)
  - Gradient Optimization of Beam Search Hyperparameters (BSc, 2019)
- Polina Kirichenko (PhD, currently a Research Scientist at Meta AI)
  - Study of Bayesian Regularization of Neural Networks (BSc, 2018)
- Nadezhda Chirkova (PhD, currently a Research Scientist at Naver Labs)
  - Bayesian Compression for Natural Language Processing, (MSc, 2018)

TEACHING

2016 - 2024	**Research seminar Machine Learning and Applications,** Faculty of Computer Science at HSE University
2015 - 2021	**Bayesian Methods in Machine Learning (organization + practical sessions),** HSE University, MSU, and Yandex School of Data Analysis
2018 - 2019	**Neurobayesian models (organization + practical sessions),** HSE University, MSU, and Yandex School of Data Analysis
2017	**Introduction to Deep Learning (Online course at Coursera, lectures)**
2016 - 2017	**Deep learning (lectures and practical sessions)**, MSU
2016 - 2017	**Machine learning (lectures)**, HSE and HSE/NES Programmes in Economics
2015 - 2017	**Data analysis (practical sessions)**, HSE University

Additionally, I was one of the main organizers and gave several lectures at the Deep|Bayes Summer School from 2017 to 2019. Together with Nadezhda Chirkova, I also led a tutorial on Bayesian machine learning at the Machine Learning in High Energy Physics Summer School (Hamburg, Germany, July 2019 and online, July 2020 and 2021), as well as at the Workshop on Machine Learning and Applications to Physics (Madrid, Spain, Dec 2019).

PROGRAM COMMITTEE

Neural Information Processing Systems, NeurIPS (reviewer):

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