Arsenii Ashukha

Home page / Google Scholar / GitHub
e-mail: ars.ashuha@gmail.com tel: +79651782915

I'm a Research Scientist at Samsung AI Center Moscow. I (almost) received a PhD in Machine Learning under supervision of Dmitry Vetrov. The results of my PhD were published at ICLR / ICML / NeurIPS and contributed to sparsification, uncertainty estimation, ensembling, computer vision, and fundamentals of Bayesian deep learning. Prior to that, I was a part of Yandex Research in collaboration with University of Amsterdam, where I worked on Bayesian deep learning with Dmitry Vetrov and Max Welling. I did ML engineering internships at Yandex (deep learning for music), Rambler (recommendation systems), and worked on NLP at Research Computing Center MSU with Natalia Loukachevitch.

PUBLICATIONS

Google Scholar: scholar.google.com/citations?user=IU-kuP8AAAAJ

* denotes joint first co-authorship

Resolution-robust Large Mask Inpainting with Fourier Convolutions, WACV 2022 — arXiv / code
Roman Suvorov, Elizaveta Logacheva, Anton Mashikhin, Anastasia Remizova, **Arsenii Ashukha,** Aleksei Silvestrov, Naejin Kong, Harshith Goka, Kiwoong Park, Victor Lempitsky

Pitfalls of In-Domain Uncertainty Estimation and Ensembling in Deep Learning, ICLR 2020 — arXiv / code
Arsenii Ashukha*, Alexander Lyzhov*, Dmitry Molchanov*, Dmitry Vetrov

Greedy Policy Search: A Simple Baseline for Learnable Test-Time Augmentation, UAI 2020 — arXiv / code
Arsenii Ashukha*, Dmitry Molchanov*, Alexander Lyzhov*, Yuliya Molchanova*, Dmitry Vetrov

The Deep Weight Prior, ICLR 2019 — arXiv / code
Arsenii Ashukha*, Andrei Atanov*, Kirill Struminsky, Dmitry Vetrov, Max Welling

Variance Networks: When Expectation Does Not Meet Your Expectations, ICLR 2019 — arXiv / code
Arsenii Ashukha*, Kirill Neklyudov*, Dmitry Molchanov*, Dmitry Vetrov

Structured Bayesian Pruning via Log-Normal Multiplicative Noise, NeurIPS 2017 — arXiv / code
Kirill Neklyudov, Dmitry Molchanov, **Arsenii Ashukha**, Dmitry Vetrov

Variational Dropout Sparsifies Deep Neural Networks, ICML 2017 — arXiv / code
Arsenii Ashukha*, Dmitry Molchanov*, Dmitry Vetrov

PROFESSIONAL EXPERIENCE

2018 April - *Now*
Research Scientist, Samsung AI Center Moscow
I'm working on the development of deep learning algorithms. Specifically, my work is focused on uncertainty estimation, robustness, and fundamentals of Bayesian deep learning. I also contributed to computer vision research e.g., image inpainting.

2016 *Feb* - 2018 April
Research Scientist, HSE & Yandex & University of Amsterdam
Created *sparse variational dropout*, *a* method for sparsification of DNNs that, for the first time, allowed to achieve over 250x compression ratio (results published at ICML'17). The modified version of SparseVD with neuron-level sparsity allowed to accelerate inference of a CNN by 2-5 times and was involved in the feature extraction for real image retrieval system (published at NeurIPS).

2016 *May* - 2016 *Sep* **Machine Learning Engineer Intern, Yandex**
I worked on feature extraction techniques for music data with convolutional neural networks. I also developed an evaluation of learned representations. The representations were used in the content-based recommendation system for Yandex music.

2015 *May* - 2015 *Oct* **Machine Learning Engineer Intern**
Worked on recommendation systems. My responsibility included improving the quality and performance of automatic feature extraction algorithms, and recommendation algorithms.

EDUCATION

2017 - 2022 **PhD** in Machine Learning, **National Research University Higher School of Economics**
Thesis title: Prior Knowledge for Deep Learning **Advisor:** Dmitry Vetrov

2015 - 2017 **MSc** in Computer Science, **Moscow Institute of Physics and Technology** (worked on sparse DNNs) with Distinction (GPA 4.87/5.0)

2011 - 2015 **BSc** in Computer Science, **Bauman Moscow State Technical University** (worked on language models)

TECHNICAL SKILLS

- **Deep Learning**, **Deep Neural Networks**, **Machine Learning**, **Modeling**;
- I'm fluent in **Python** and I used to code in C/C++, Go, language is not a problem after all.
- I'm also fluent with common data science tools such as **NumPy**, **matplotlib**, **scikit-learn**, **pandas**.
- I'm comfortable with the common data science environment e.g., **bash**, **git**, **Linux**.
- My primary deep learning framework is **PyTorch**. Prior to that, I had an experience with Theano+Lasagne and TensorFlow.
- Comfortable with **GPU clusters and distributed training**.
- I have experience with **MapReduce**, Hadoop, Hive, and Spark.

CODE

- **Simple MVP implementations of ML algorithms:**
  - Real NVP normalizing flows
  - Quantile Regression DQN (Distributional RL)
  - Equivariant GNN
- **Research-ready implementations:**
  - Multi-GPU SimCLRv1 closely reproduced results on both CIFAR-10 and ImageNet
  - Ensembles (Deep ensembles, Snapshot ensembles, cSGLD, FGE, etc.)

THESIS CO-SUPERVISON

- Alexander Lyzhov (moved to NYU)
  - Deep Neural Network Ensembles: Analysis and Approaches to Diversification (MSc, 2020)
- Andrei Atanov (moved to EPFL)
  - Effective Learning of Deep Neural Networks Ensembles (BSc, 2018)
  - Learning Deep Models with Small Data (MSc, 2020)
- Evgenii Nikishin (moved to Mila)
  - Stability Improvement and Knowledge Transfer in Deep Reinforcement Learning (MSc, 2019)

PROGRAM COMMITTEE

- Neural Information Processing Systems, NeurIPS:
  - 2019: top-50% highest-scored reviewers
  - **2021: outstanding reviewer award (top-8%)**
- International Conference on Machine Learning, ICML (2019, 2020)
  - 2020: top-33% highest-scored reviewers
- International Conference on Learning Representations, ICLR (2020, 2021)
- ICML Workshop on Invertible Neural Networks (2019, 2021, invertibleworkshop.github.io)
- Bayesian Deep Learning Workshop (since 2017, bayesiandeeplearning.org)

TEACHING

- Supervisor of reading clubs on machine learning at HSE and Yandex school of data analysis (since 2017)
- A lecture with a practical session on Normalizing Flowsat DeepBayes Summer School (since 2017)
- Lecturer, Moscow Institute of Physics and Technology: I was a lecturer and manager of the deep learning brunch of a facility-wide machine learning course ~60 students (ml-mipt.github.io). Also, I taught deep learning and practical sessions on cutting-edge ML algorithms on a facultative course "Data Mining in Action" ~ 200 students (bit.ly/3eRLGYp). The goal of this course is to make ML education available for everyone for free.