Konstantin Mishchenko

konstantin.mishchenko@kaust.edu.sa
Last updated: April 11, 2019

Education

2017-2020 KAUST, PhD in Computer Science, Adviser: Peter Richtárik
2016–2017 ENS Cachan and Paris-Dauphine, MSc in Machine Learning
2012–2016 Moscow Institute of Physics and Technology BSc in Computer Science and Physics

Conferences

2019 International Conference on Machine Learning, Program Committee Member
2018 Conference on Neural Information Processing Systems, Montreal, Poster
2018 International Conference on Machine Learning, Stockholm, Oral
2018 International Symposium on Mathematical Programming, Bordeaux, Invited talk
2018 Informs Optimization Society Meeting, Denver, Invited Talk

Internships and projects

2018 Applied Scientist Intern at Amazon
2017 Pre-Doc Summer School, ETH Zurich
2017 Machine Learning Summer School (Acceptance rate 15%, the only accepted undergrad student), MPI Tübingen
2017 Google Machine Learning Summit, Zurich, Poster
2017 Course project on air pollution prediction for Plume Labs data challenge, **1st place** in the competition
2016 Course project on using dropouts for learning restricted Boltzman machines
2016 C++ development summer intern, AIM High Tech (High Frequency Trading), Moscow

Achievements and awards

2018 IEEEXtreme team programming competition, 71st place worldwide
2017 IEEEXtreme team programming competition, 123rd place worldwide
2016-2017 Paris Graduate School of Mathematics fellowship (highly competitive scholarship with about 20 laureates every year)
2015 Higher School of Economics Olympiad on Applied Math and Informatics, 1st prize
2014 Abramov-Frolov scholarship for excellence in study
2012 Top-1 (max score) at the National Exam in math (only 54 participants out of >800k scored max)
2012 Moscow Mathematical Olympiad, 1st prize
2012 Phystech International, Mathematics, 1st prize

Research interests

- Stochastic and Distributed Algorithms
- Nonconvex Optimization
- Machine learning

Papers

- Stochastic Distributed Learning with Gradient Quantization and Variance Reduction
- 99% of Parallel Optimization is Inevitably a Waste of Time
- Distributed Learning with Compressed Gradient Differences
- SEGA: Variance Reduction via Gradient Sketching, Published in proceeding of the Conference on Neural Information Processing Systems (**NeurIPS, Poster**), 2018
- A Stochastic Penalty Model for Convex and Nonconvex Optimization with Big Constraints
- A Distributed Flexible Delay-tolerant Proximal Gradient Algorithm
- A Delay-Tolerant Proximal-Gradient Algorithm for Distributed Learning, Published in proceedings of the International Conference on Machine Learning (**ICML, Oral**), 2018.

Languages

Russian	Native
English	Fluent
French	Intermediate