Ekaterina Antonenko

Education

since 2020 **École Polytechnique**, *PhD candidate*.
Data Science and Mining (DaSciM) team, Laboratoire d'informatique (LIX), CIFRE with Digitalent
Scientific advisor: Jesse Read

2014–2016 **British Higher School of Art & Design**, *2D-animation*.

2012–2014 **Moscow Bioinformatics School**.

2009–2014 **Moscow State University**, *Diploma in Mathematics*.

Employment

since 2020 **Data scientist**, *Digitalent (Paris, France)*.

2017–2019 **Math instructor**, *Russian School of Mathematics (New York, USA)*.

2013–2016 **Analyst**, *Moscow Center for Continuous Mathematical Education (Moscow, Russia)*.

Preprints

1. E. Antonenko, R. Beigaitė, M. Mechenich, J. Read and I. Žliobaitė, *Backward inference in probabilistic Regressor Chains with distributional constraints*.
2. E. Antonenko, A.Carreño, J. Read, *Autoreplicative Random Forests for missing value imputation*.
3. M. Konnova, E. Antonenko, J. Read, *Missing value imputation for genomics data using a Sequence Based Generative Adversarial Network (SBGAN)*.

Publications

1. E. Antonenko, J. Read, *Chains of Autoreplicative Random Forests for missing value imputation in high-dimensional datasets*, [Best paper award], Multi-Label Learning workshop at the ECML conference, 2022, `https://mll2022.csd.auth.gr/papers/ChARF%20camera-ready.pdf` .
2. E. Antonenko, J. Read, *Multi-modal ensembles of regressor chains for multi-output prediction*, Advances in Intelligent Data Analysis XXI - 21st International Symposium, IDA 2022, `https://link.springer.com/chapter/10.1007/978-3-031-01333-1_1`.
3. V. Ivanenko, E. Antonenko, M. Gelfand, J. Yager, F. Ferrari, *Changes in segmentation and setation along the anterior/posterior axis of the homonomous trunk limbs of a remipede (Crustacea, Arthropoda)*, PeerJ, 2016, `https://peerj.com/articles/2305/`.

Talks

1. *Autoreplicative Random Forests for missing value imputation*, Group seminar, KU Leuven KULAK, Kortrijk, Belgium, February 2023.
2. [Poster] *Genotype imputation with multi-label Random Forests*, Machine Learning in Computational Biology, online, November 2022.

3. *Chains of Autoreplicative Random Forests for missing value imputation in high-dimensional datasets*, Multi-Label Learning workshop: current trends and open challenges, ECML PKDD 2022, Grenoble, France, September 2022.
4. *Multi-modal ensembles of Regressor Chains for multi-output prediction*, Intelligent Data Analysis XXI - 21st International Symposium, IDA 2022, Rennes, France, April 2022.

Teaching

- CSE204 Machine Learning (Bachelor Programme), Teaching assistant, *École Polytechnique*, Spring 2022, 2023.

Mentoring students

- Margarita Konnova (École Polytechnique), Bachelor thesis *"Missing Value Imputation for Genomics Data using a Sequence Based Generative Adversarial Network"*, coadvised with Jesse Read, 2022.

Additional training

- Data Science: Multiple Imputation in Practice, *Utrecht University, The Netherlands*, 2022.
- Introduction to quantitative genetics, *MIPT, Russia*, 2021.
- Summer School on Machine Learning in Bioinformatics, *HSE University, Russia*, 2020.

Completed online-courses

- AI for Medicine Specialization, *Deeplearning.ai*, 2021, *Certificate*.
- Deep Learning Specialization, *Deeplearning.ai*, 2020, *Certificate*.
- Machine Learning, *Stanford University*, 2019, *Certificate*.
- Molecular Biology – Part 3: RNA Processing and Translation, *Massachusetts Institute of Technology*, 2019, *Certificate*.
- Molecular Biology – Part 2: Transcription and Transposition, *Massachusetts Institute of Technology*, 2019, *Certificate*.
- Molecular Biology – Part 1: DNA Replication and Repair, *Massachusetts Institute of Technology*, 2018, *Certificate*.
- Quantitative Biology Workshop, *Massachusetts Institute of Technology*, 2018, *Certificate*.
- Introduction to Biology – The Secret of Life, *Massachusetts Institute of Technology*, 2018, *Certificate*.
- Bioinformatics Algorithms (Part 1), *University of California, San Diego*, 2014, *Certificate*.

Additional skills

- Programming: Python (sklearn, tensorflow, keras), R, MatLab
- Languages: English (fluent), French (intermediate), Russian (native)