

DARIA RYABCHENKO

github.com/daryabchenko | podariya12@gmail.com

EDUCATION

Skolkovo Institute of Science and Technology

PhD in Computational and Data Science and Engineering

Molecular generation and structure-based design; focus on GNNs, diffusion, and flow matching.

Sep.2024-Present

Moscow, Russia

National Research University Higher School of Economics

MSc in Data Science. GPA: 8.0/10.0.

GNNs and generative modeling; focus on Transformer models.

Sep.2022-May 2024

Moscow, Russia

Ufa State Petroleum Technological University

BSc in Biotechnology. GPA: 3.92/4.0; ranked 1st in class.

Sep.2017-May 2021

Ufa, Russia



WORK EXPERIENCE

Ligand Pro

ML Specialist

MOLECULAR GENERATION: design and run reproducible ML experiments, compare model architectures and training strategies, analyze performance metrics, and identify changes that improve model quality and training stability.

BENCHMARKING: developed **HEDGEHOG**, an automated benchmarking pipeline for generative ML models, including data preprocessing, structural validation, metric computation, and comparative model evaluation.

ML TOOLING: built an MVP for medicinal chemists to interact with molecular generation models and inspect generated molecules; implemented and tested methods from recent research papers.

Sep.2024-Present

Moscow, Russia

One Cell

Data Analyst

Coordinated data preparation and annotation workflows, translating ML requirements into data collection and labeling tasks. Analyzed and identified gaps in datasets. Worked with pathologists and ML engineers to ensure data quality for modeling.

Jun.2024-Nov.2024

Moscow, Russia

TECHNICAL SKILLS

Programming: Python, SQL, Go, Bash, Jupyter.

Machine Learning: classification, regression, clustering, feature engineering, cross-validation, hyperparameter optimization.

Deep Learning: Graph Neural Networks (GNNs), Transformers, Diffusion Models, Flow Matching, Autoregressive Models, Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs).

ML & Scientific Computing: PyTorch, PyTorch Geometric, scikit-learn, NumPy, Pandas.

Cheminformatics & Molecular Modeling: RDKit, Open Babel, Schrödinger Suite.

Development & Experimentation: Git, Docker, ClearML, Weights & Biases.

AI-assisted Workflows: ChatGPT, Codex, LLM-based coding agents.

SELECTED CONFERENCE PRESENTATIONS

BCADD 2025 – XXXI Symposium on Bioinformatics and Computer-Aided Drug Discovery

Oct.2025

Do-No-Harm Molecular Generation: 12-model Benchmark and KRAS G12D Case Study, p. 58.

D. Ryabchenko et al. – Oral presentation.

AIChem 2025 – II Conference on Artificial Intelligence in Chemistry and Materials Science

Nov.2025

Evaluation of the Reliability of Molecular Generation Models: A KRAS Case Study, p. 128.

D. Ryabchenko et al. – Oral presentation. **Best Oral Presentation Award.**

Sirius.Biotech 2026

May 2026

HEDGEHOG: A System for Evaluating Molecular Generators through Rigorous Filtration, p. 14.

D. Ryabchenko et al. – Poster presentation.

ADDITIONAL INFORMATION

Scientific Interests: machine learning, deep learning, graph machine learning, machine learning for chemistry and biology, molecular generation, generative modeling, computer-aided drug design.

Background: mathematics, biotechnology, chemistry, and biology.

Languages: Russian (Native), English (C1), German (B1).