Vasily Ilin

Google Scholar: Vasily Ilin | Github: Vilin97 | email: vasilin97@gmail.com

Education

University of Washington | 2020 - 2026
Applied Mathematics, PhD

Boston University | 2015 - 2020
Computer Science, MS
Mathematics, BA & MA (Magna Cum Laude). Thesis: "Stochastic Simulation Algorithms and Benchmarks", paper.

Skills

Technical Skills: Python, Jax, SQL, Julia, Java, Git, Lean
Math Skills: Deep learning, diffusion modeling, sampling, optimal transport, mean-field limits, SDE & PDE, math formalization

Publications and Talks

Community and Mentorship Through the Experimental Lean Lab (talk at JMM) | 2025 | abstract
We share lessons we've learned in building community and mentoring undergraduate research projects in Lean.

RealEdit: Reddit Edits As a Large-scale Empirical Dataset for Image Transformations (submitted to CVPR) | 2025
A dataset and diffusion generative model to perform arbitrary image edits. The first global edit model trained on real data.

Deterministic sampling with adaptive score based transport modeling (first author, talk at JMM) | 2025 | abstract
An algorithm for deterministic sampling by integrating the Fokker-Planck equation using particles.

Transport based particle methods for the Fokker-Planck-Landau equation (first author, submitted to CMS) | 2024 | pre-print
An algorithm for simulation of plasma using a neural network, inspired by score-based generative modeling.

Extending `JumpProcesses.jl` for fast point process simulation with time-varying intensities | 2024 | paper
An algorithm to efficiently simulate any point process on the real line with a continuous intensity rate.

Catalyst: fast biochemical modeling with Julia (PLOS Comp Bio) | 2023 | paper
Julia library for modeling and high-performance simulation of chemical reaction networks.

Work Experience

ML Intern at Google Cloud (Python, SQL) | Summer 2024
Trained XGBoost to predict defective TPUs from HBM ECC telemetry, achieved 80% accuracy and improved test efficiency 4x.

ML Intern at YouTube (Python, C++, SQL) | Summer 2023
Trained ML models to enhance cache efficiency and achieved a 2x Egress/Ingress improvement.
Proposed and implemented a better load balancing algorithm based on the two random choices model.

Data Engineering Intern at Google (FlumeJava, SQL) | Summer 2022
Built an end-to-end distributed pipeline for Nearby Share, from Android logs to SQL tables with statistical metrics.
Brought dashboard loading time down from 30s to 1s using approximate aggregation.

Google Summer of Code (Julia) | Summer 2021 | code, blog post
Implemented, tested, benchmarked, and optimized algorithms to simulate jump processes.
Improved run-time of space simulations by 70%, resulting in two publications.

Projects

AI DnD Bot (OpenAI API, PostgreSQL) | Summer 2024 | code
Made a Telegram bot for Dungeons & Dragons, including short- and long-term memory, image generation, and persistence.

Deep Reinforcement Learning in DeepRacer | Spring 2022 | video, report, code
Trained 80+ different models, placed in top 3% out of 1000+ participants in a virtual race hosted by AWS.

Leadership

Mathematical Formalization (Lean) | 2021-present | code 1, 2, 3, 4, 5, 6, 7, 8, 9
Developed and taught a new class on mathematical formalization. Started the Math AI seminar with speakers from academia and industry. Taught groups of 10+ undergrads. Contributed to mathlib (1, 2, 3). Featured in JMM 2025 (link).