Timofei Rusalev

+7 (985) 316-18-90 ⋄ Moscow, Russia
rusalev.timofey@gmail.com ⋄ GitHub ⋄ arXiv

Seeking a ML Research Engineer role to apply 5+ years of theoretical physics research and AI research internship experience toward R&D challenges, focusing on large language models.

Education

PhD Candidate in Mathematical Physics Oct 2021 – Oct 2025
Steklov Mathematical Institute, Moscow, Russia
Dissertation defense expected in 2026

Master of Science in Physics, with honors Sep 2019 – Jun 2021
Lomonosov Moscow State University, Moscow, Russia

Bachelor of Science in Physics, with honors Sep 2015 – Jun 2019
Lomonosov Moscow State University, Moscow, Russia

Experience

Research Intern Aug 2025 – Present
T-bank, AI research development, Moscow, Russia

- Studied cross-layer and cross-seed matching of GPT-2 Small SAE features, characterizing alignment quality via explained variance.

Research Intern Apr 2021 – 2025
Steklov Mathematical Institute, Moscow, Russia

- Conducted scientific research in theoretical and mathematical physics, resulting in 5 publications — 3 of which appeared in leading international Q1 journals.
- Engineered a novel framework for dimensionality reduction of Rényi entropy computations, mapping a complex high-dimensional problem to a tractable low-dimensional equivalent.
- Derived closed-form analytical solutions for Rényi entropies by combining the replica trick with advanced saddle-point approximations for gravitational path integrals.

ML Projects

Cluster Evolution of Sparse Autoencoders Features Across Layers (GitHub)

- Built a research pipeline for sparse autoencoders (SAE v5-32k) across all 12 layers of GPT-2 small, applying UMAP + HDBSCAN clustering to both decoder columns and feature text embeddings (generated via SBERT all-mpnet-base-v2 on Neuronpedia descriptions).
- Applied two cluster-evolution methods, i) SAE Match for individual feature alignments and ii) optimal transport, to evaluate concept evolution in both decoder-column and text-embedding spaces.
- **Key findings:** i) three-phase concept evolution (Fragmentation → Consolidation → Specialization), ii) text-description embeddings outperform decoder columns , iii) optimal transport outperforms SAE Match.

Skills

- **Programming:** Python (NumPy, SciPy, Pandas, PyTorch, scikit-learn, Matplotlib, Seaborn), Wolfram Mathematica, Jupyter Notebook, Git
- **Machine Learning:** linear models, decision trees, ensembles, gradient boosting, neural networks (MLP, RNN, transformers), NLP, LLM, VLM, clustering, dimensionality reduction, mechanistic interpretability