

Ekaterina Antipushina

✉ Ekaterina.Antipushina@skoltech.ru

☎ +79165741784

🌐 <https://www.linkedin.com/in/katherine-antipushina/>

🔗 <https://github.com/utoprey>

Profile

DL-research in medical AI skilled in programming Python. Having experience in Pandas, NumPy, Seaborn, Scikit-learn, Pytorch. My research interests center on quantization, foundation model development, multimodality, data fusion approaches and generative AI (VAEs, GANs, NOT).

Skills & Proficiencies

Software & Tools

Python (Visualisation/EDA/ETL/PyTorch), git, Docker, LaTeX, HuggingFace, Transformers

Languages

Russian (native), English (C1 Advanced)

Work Experience

2024/06 – 2025/06
Moscow, Russia

ML RESEARCH ENGINEER

Applied AI Center

Project 1. EEG Foundation Models & Neural Quantization

Developed quantization techniques for converting continuous 1D data (EEG) into discrete token representations.

Architected foundation model for multichannel 1D data with specialized temporal encoding. Investigated cross-modal learning approaches for 1D to 4D integration and multimodal neural signal processing.

Stack: PyTorch, transformers, vector quantization, self-supervised learning, neural tokenization, foundation models

Project 2. Pharmacological Biomarker Discovery & Neuron-Glial Interaction Modeling in Brain Pathologies

- Formulated hypotheses based on experimental conditions and biological processes, validated through statistical testing.
- Conducted exploratory data analysis (EDA) to identify key trends, guiding feature selection.
- Performed feature selection to reduce dimensionality and enhance model interpretability.
- Used clustering techniques to segment data by cell behavior and treatment responses.
- Built and evaluated machine learning models, applying cross-validation for performance assessment.
- Integrated statistical analysis and machine learning methods in a complete pipeline for model development and evaluation.

Stack: Python, Pandas, NumPy, SciPy, Scikit-learn, statistical hypothesis testing, PCA/feature selection, K-means clustering, cross-validation, supervised/unsupervised ML models

2024/06 – 2024/12
Moscow, Russia

RESEARCHER

Applied AI Center, Neuroimaging and Cognitive Neuroscience Laboratory 📍

Project 1. pyOpenNFT: Open-Source Framework for Real-Time fMRI Neurofeedback Training.

- Designed and implemented fully Python-based architecture replacing MATLAB dependencies, featuring parallelized workflows with shared memory buffers for faster volume processing.
- Developed FastAPI-based prediction server with RESTful interface and Lab Streaming Layer integration for real-time EEG-to-fMRI signal prediction.

- Implemented real-time quality assessment module and optimized numerical computations using NumPy+MKL, achieving cross-platform compatibility on Windows and Linux systems

Stack: Python, FastAPI, NumPy, SciPy, nibabel, PyQt6, OpenCV, Lab Streaming Layer (LSL), REST APIs, Poetry, cross-platform deployment (Windows/Linux)

Project 2. Computational Framework for Psychophysiological State Monitoring During Emotion Regulation (Samara University Partnership)

- Designed and implemented data processing pipelines for analyzing and visualizing physiological signals, and EEG, collected during controlled experiments
- Developed end-to-end data workflows for signal preprocessing, feature extraction, and visualization, ensuring scalability and reproducibility across multiple experimental conditions
- Conducted statistical analysis on test results to evaluate participant performance under various task conditions, utilizing tools such as Python for hypothesis testing and performance metrics
- Built and optimized machine learning models to analyze physiological responses, employing techniques such as feature selection, classification

2023/09 – 2023/12
Moscow, Russia

JUNIOR ML RESEARCH ENGINEER

Applied AI Center, Center of Excellence in Precision Medicine at the University of Sharjah

Project 1. Rest2Task: Image-to-Image model for task-based fMRI prediction

Developed a generative AI framework to create task-based fMRI data from resting-state scans

- **Techniques and Models:** Variational Autoencoders (VAE), Conditional GAN (cGAN), Wasserstein Conditional GAN with Quadratic Cost (WCGAN-QC), Vanilla Neural Optimal Transport (NOT), Deformable NOT
- **Outcome:** Configured Docker and integrated Wandb for experiment tracking. Generated task-based fMRI data from resting-state scans (MSE: 0.05 ± 0.01)

Project 2. Multimodal Neuroimaging Analysis for Schizophrenia Prediction

Development of an end-to-end schizophrenia prediction tool based on multimodal neuroimaging data

- **Techniques:** Early/Late Fusion, Feature Selection (PCA, Local Linear Embeddings, LDA, Isomap, t-SNE)

Project 2.1. Geometric Machine Learning for Schizophrenia Prediction

Domain: fMRI, Topological Data Analysis (TDA)

Development of a schizophrenia prediction model using geometric features

- **Techniques:** Topological Data Analysis (Persistent Homology), Manifold Learning
- **Outcome:** Enhanced schizophrenia prediction models using geometric and topological neuroimaging features, including Betti numbers, spectral properties, and graph representations

Project 3: ML in Neuropsychiatry and Oncology

Applied ML to analyze molecular mechanisms and identify diagnostic/prognostic biomarkers for depressive disorders, as well as blood and colorectal cancers.

- **Outcome:** Developed predictive models for disease biomarkers, utilizing omics data and clinical features to improve diagnostic and prognostic accuracy.

2023
Moscow, Russia

INTERN DATA ENGINEER

AIRI 


- Conducted exploratory data analysis (EDA) on neuroimaging data to identify meaningful patterns and correlations.
- Utilized Python (NumPy, Pandas) for preprocessing and integrating neuroimaging datasets.

- Developed and fine-tuned machine learning models to analyze fMRI and DTI data.
- Applied feature selection techniques to identify the most relevant features for predictive modeling.
- Implemented data fusion (early, late, intermediate) methods to combine neuroimaging data types.
- Designed and evaluated machine learning workflows using scikit-learn and Wandb or TensorFlow

2022

Moscow, Russia

RESEARCH ASSISTANT

Skoltech, Applied AI center, NeuroML group 

The focus of the tasks is in the development and implementation of machine and deep learning models for the analysis of neuroimaging data including fMRI, sMRI, EEG. Collecting and organizing datasets, preprocessing and analyzing data using various software tools (MNE, Nilearn).

Projects

2025

Improving speech through knowledge distillation using WavLM model

- Developed a PyTorch system for speech noise suppression, where a student model is trained on noisy data, simulating the embeddings of a teacher's WavLM model trained on clean speech.
- Used the VoiceBank+DEMAND dataset with 10 types of noise at SNR 0-15 dB, applying a combined loss function (cosine similarity + MSE), which reduced the error from 0.184 to 0.074.

2024

Multimodal Models with Narrow Domain

AIRI

This project focuses on fine-tuning multimodal models to enhance their performance in specialized, narrow-domain applications

Models: Bridge Matching, Stable Diffusion 1.5, GILL Model, TinyLLaVA-Phi-2-SigLIP-3.1B

Education

2024 – present


PhD-1 in Computational and Data Science and Engineering

Skoltech, Applied AI

2022 – 2024

Moscow, Russia

MASTER'S DEGREE - Data Science

Biomedically-Informed Artificial Intelligence Laboratory, Skoltech 

Master's Thesis: "Interpretable machine learning models for the analysis of multimodal neuroimaging and biomedical data using data fusion methods in the diagnosis of schizophrenia", Supervisor: Candidate of Physical and Mathematical Sciences, Sharaev M.

2018 – 2022

Moscow, Russia

BACHELOR'S DEGREE - Biomedical Engineering Systems & Technologies

Moscow Aviation Institute

Courses

2024

Moscow, Russia

Professional development certificate in foundations of Generative Artificial Intelligence

Skolokovo Institute of Science and Technology

2023

Moscow, Russia

Professional development certificate in the use of modern neural networks for natural language processing tasks

Skolkovo Institute of Science and Technology

2023/08 Altay	Summer School of Machine Learning at Skoltech <i>Skoltech</i> Worked on project dedicated to geometric machine learning for schizophrenia prediction. In this work, topological features (betty numbers, graph and spectral features) were added to standard features extracted from neuroimaging data and topological invariants were calculated.
2022 Tomsk, Russia	Professional development certificate of IT technology in Healthcare <i>Tomsk Polytechnic University</i> Course program included seminars on different types of data science and machine learning algorithms, discussion sessions on how to use data-driven models to solve real-world problems, popular tools and libraries for data science.

Publications

2025	Antipushina E., et al. pyOpenNFT: an open-source Python framework for ML-based real-time fMRI and EEG-fMRI neurofeedback ↗ <i>A* conference MICCAI 2025, Lecture Notes in Computer Science, Springer</i>
2025	Antipushina E., et al. CSTNet: A Generative Framework for EEG-to-ECoG Translation via Optimal Transport <i>A* conference MICCAI 2025, Lecture Notes in Computer Science, Springer</i>
2024	Maria Boyko, Ekaterina Antipushina, et al. Interpretable AI models for predicting distant metastasis development based on genetic data: kidney cancer example ↗ <i>BIO Web of Conferences</i>
2023	Piliugin N., Knyshenko M., Sain A., Antipushina E. Artificial Sensations Evoked by Transcutaneous Electrical Nerve Stimulation: Investigating the Parameters Space ↗ <i>IEEE Ural-Siberian Conference on Computational Technologies in Cognitive Science, Genomics and Biomedicine (CSGB)</i>
2020	Antipushina E.S. Immersive virtual reality systems in medical care. Collection on the results of the International Scientific Conference «Scientific research of the SCO countries: synergy and integration». DOI 10.34660/INF.2020.89.47.013. (Eng)
2020	Kolosov A.S., Antipushina E.S. Assisted circulation as a method of treating heart failure and as a way to support donor organs. Collection of materials of the XXXIII All-Russian scientific and technical conference of students, young scientists and specialists "Biotechnical, medical and ecological systems, measuring devices and robotic complexes (BIOMEDSYSTEMS - 2020). (Rus)