



Samatov Denis

MACHINE LEARNING ENGINEER /
DATA SCIENTIST

ABOUT ME

I am a Data Science and Machine Learning specialist with 4+ years of experience and a solid academic background in Applied Mathematics and Computer Science. My expertise lies in developing applied AI systems, particularly in medical diagnostics, image segmentation, radiomics, and scientific computing. I have end-to-end experience across the machine learning pipeline and a strong focus on combining algorithmic methods with domain knowledge to deliver reliable and impactful solutions in both independent and team settings.

TOPICS OF QUALIFICATION WORKS

- Automatic Segmentation of Epicardial Fat and Quantitative Assessment of Radiomic Parameters in Cardiac Computed Tomography
- Machine Learning Methods in Radiomics for Analyzing Cardiovascular Images
- Recognition of Russian-Language Text in Financial Reports Using Neural Networks
- Beam Parameter Restoration at the NICA Accelerator Complex Based on Beam Position Monitor Data
- Option Strategies and Practical Applications of Options
- Tensor-Based Modal Decomposition for Reduced-Order Modeling of High-Dimensional Spatiotemporal Systems
- Optimal Sparse Sensor Placement via Tensor QR Decomposition with Column Selection in High-Dimensional Forecasting Tasks

PUBLICATIONS

- Samatov D.S. Methods of Machine Learning in Radiomics for Cardiovascular Image Analysis. // Modern Problems of Mathematics and Its Applications: Proceedings of the International (54th All-Russian) Youth School-Conference. — February 2023
- Samatov D.S., Mochula A.V. Methods of Machine Learning in Radiomics for Cardiovascular Image Analysis. // Perspectives of Fundamental Sciences Development: Proceedings of the International XX Conference of Students, Postgraduates, and Young Scientists. — April 2023

EDUCATION

Tomsk Polytechnic University

- Master's Degree, Applied Mathematics and Computer Science 2024 – present
- Bachelor's Degree, Applied Mathematics and Computer Science 2020 – 2024

ML / DL TOOLS & FRAMEWORKS

- Python, NumPy, Pandas, Optuna, OpenCV, TensorLy, Radiomics, Matplotlib, Seaborn, Plotly
- Classical ML: scikit-learn, XGBoost, LightGBM, CatBoost
- DL: PyTorch, TensorFlow, Hugging Face Transformers
- LLM & RAG: LangChain, LlamaIndex, SentenceTransformers
- MLOps: MLflow, Weights & Biases, Hydra, Docker
- Deployment: FastAPI, Flask, Docker, TensorRT
- Working with Unbalanced Data
- Interpretation of models: SHAP, LIME

WORK EXPERIENCE

TUSUR | LPMITF | 06.2023 - 12.2023

- Participated in a project to develop recommendation systems for the Cardiology Research Institute using machine learning methods in radiology.
- Responsible for the collection, cleaning, and analysis of medical data, ensuring its accuracy and reliability for subsequent analysis. Reduced processing time by 30% through data cleaning process automation.
- Developed predictive models and recommendation algorithms based on processed data using various machine learning methods. Improved prediction accuracy by 25% compared to previous models.

Fintech Association | 08.2023 - 11.2023

- Actively participated in the development of a prompt generator presented at Finopolis 2023.
- My primary responsibility was compiling a Russian-language financial dataset and selecting optimal machine learning models for the project.
- Reduced data preparation and cleaning time by 25% through automation of information collection and processing.
- Improved prompt generation accuracy by 30% compared to previous models using YandexGPT and Gigachat APIs.
- Decreased false positives in the bank offer selection algorithm by 20% through precise model architecture selection and training ruGPT 3.5 models using LoRA.
- Conducted testing of various machine learning models to evaluate their performance and suitability for integration into the prompt generator.

- Samatov Denis, Shandov Mikhail. Beam Parameters Restoration at the NICA Accelerator Complex – Based on the Beam Position Monitor Data. // START, Joint Institute for Nuclear Research, Dubna, 2023
- Samatov D.S. Automatic Segmentation of Epicardial Fat and Quantitative Evaluation of Radiomic Parameters in Cardiac Computed Tomography. // Perspectives of Fundamental Sciences Development: Proceedings of the International XXI Conference of Students, Postgraduates, and Young Scientists. — April 2024
- Maksimova A.S., Samatov D.S., et al. Capabilities of Radiomic Analysis of Cardiac MRI Images in Cine Mode for Identifying Post-Infarction Areas of the Left Ventricular Myocardium. // Digital Diagnostics.
- Zavadovsky K.V., Samatov D.S., Merzlikin B.S., et al. EPIFAT – Module for Automatic Segmentation of Epicardial Adipose Tissue on Cardiac CT Images // Certificate of State Registration of Computer Program No. 2025610317 dated 09.01.2025, Tomsk National Research Medical Center of the Russian Academy of Sciences. Registered by the Federal Service for Intellectual Property (Rospatent).

ACHIEVEMENTS

- Special Nomination Winner of the National Technology Olympiad (Student Track) in "Computer Vision Technologies and Digital Services"
- Prize Winner of the Hackathon at the MIPT Educational Forum in Artificial Intelligence, Mathematics, and Physics
- Winner of the "Artificial Intelligence and Machine Learning" Track at the FINOdays Hackathon
- Second-degree Diploma for the Presentation titled "Automatic Segmentation of Epicardial Fat and Quantitative Assessment of Radiomic Parameters in Cardiac Computed Tomography" at the "Prospects for the Development of Fundamental Sciences: Proceedings of the XXI International Conference for Students, Postgraduates, and Young Scientists," held in April 2024
- Participant in the ASCOMP School, Innopolis, October 2024.
- Participant in the ICOMP Conference (International Conference on Computational Optimization), Innopolis, 2024.
- Presenter at the Congress of the Russian Society of Radiologists and Radiologists, Moscow, November 2024. Presented a report titled "Development of Methods for Segmentation of Epicardial Adipose Tissue Using Artificial Intelligence".
- Participant in the School-Conference on Tensor Methods in Mathematics and Problems of Artificial Intelligence, China, November 2024.

Joint Institute for Nuclear Research (JINR) | 06.2023 - 08.2023

- Participated in a project to develop a particle track reconstruction algorithm for the accelerator.
- Optimized the particle track reconstruction algorithm to increase processing speed and improve track reconstruction quality. This optimization resulted in a 25% increase in the algorithm's speed.
- Developed machine learning algorithms to enhance the analysis of multiple particle tracks in the beam, improving the efficiency and accuracy of the reconstruction process by 12%.

Cardiology Research Institute | 03.2021 - 05.2023

- Developed an algorithm for automatic segmentation of cardiac CT and MRI images.
- Created a comprehensive cardiac medical image processing algorithm, from preprocessing and segmentation to ROI extraction and post-processing, utilizing statistical and mathematical image analysis methods. This algorithm increased image processing speed by up to 40%, while maintaining segmentation accuracy at 70%.
- Compiled a dataset for training neural networks for segmentation.
- Trained Unet / Attention UNet models for medical image segmentation, achieving a segmentation accuracy of over 80%.
- Applied texture analysis methods to identify pathologies using radiomics, enhancing diagnostic accuracy by 15% and improving early-stage hidden pathology detection.
- Developed an application with two automatic segmentation methods and manual ROI adjustment, increasing the efficiency of physicians' workflows.

Analytics and Machine Learning Department at MSUU | 04.2024 - now

- Developed and implemented an advanced internal RAG system to enhance document retrieval efficiency and accuracy within the organization.
- Collaborated with the team to develop a machine learning-based salary monitoring system for the recruitment department, enabling targeted salary recommendations for specific vacancies and optimizing candidate selection.
- Implemented a multimodal AI-driven evaluation system for automated candidate selection using video content, significantly increasing recruitment efficiency and analytical accuracy.

Heriot-Watt TPU Center | 09.2024 - now

- Developing tensor-based algorithms to reduce complexity in large spatiotemporal datasets, enabling efficient modeling and analysis of dynamic systems.
- Utilizing tensor-based QR decomposition methods to strategically select sensor locations, improving forecasting accuracy and reducing data dimensionality in high-dimensional systems.
- Integrating reinforcement learning methods into geological problem-solving to enhance decision-making processes, optimize resource exploration, and manage uncertainty in geological systems.