

Pavel Repnikov

Location: Moscow, Russia

GitHub | Email: prepnik@mail.ru | Mobile: +79888639559

TECHNICAL SKILLS

Languages	: C++, Python, SQL
Frameworks	: Unreal Engine, Unity, Pytorch
Libraries	: numpy, pandas, PySpark, Greenplum, XGBoost, LightGBM, CatBoost, numba, sklearn, geopandas, SciPy, PySR, PyClustering, Optuna




WORK EXPERIENCE

Data Scientist Intern <i>Sberbank, Risk Modeling Department</i>	Oct 2023 – Present <i>Moscow, Russia</i>
<ul style="list-style-type: none">• A model for determining the high probability of a client entering bankruptcy status• A model for detecting fraud in consumer loans<ul style="list-style-type: none">* the full cycle of model development* terms of reference, data collection* model development* communication with the customer of the task* preparing the model for production• Geolocation data. Geolocation data in fraud prevention• Behavior modeling. Probabilities of events as a new way of explaining incidents• Time series forecasting• Working with classic tabular data• Building datasets from multiple data sources	

EDUCATION

Lomonosov Moscow State University <i>MSc in Physics, Chair of Mathematical Modeling and Computer Science</i>	2022-2024 <i>Moscow, Russia</i>
Lomonosov Moscow State University <i>BSc in Physics, Chair of Mathematical Modeling and Computer Science</i>	2018-2022 <i>Moscow, Russia</i>

COMPLETED PROJECTS

<u>Forecasting global population dynamics</u>	<i>Python, numba, geopandas</i>	 code
<ul style="list-style-type: none">• Partial differential equations as a way to predict the population on the globe• Modification of the classical formulation of the problem taking into account spatial components• A solution on a set of arbitrary shape		
<u>Credit Scoring on a synthetic dataset</u>	<i>Python, XGBoost, LightGBM, CatBoost, PyClustering, SciPy</i>	 code
<ul style="list-style-type: none">• The divide and conquer principle. Building models independently for different years• Automatic feature generation• Testing statistical hypotheses• Clustering of tabular data		
<u>Furniture object detection</u>	<i>SQL, C#, Unity, Python, Pytorch</i>	 code
<ul style="list-style-type: none">• Creating a synthetic dataset using Unity• Object detection finetuning using Pytorch		

Bayesian Decision Making as a Theoretical Basis for a New Look at Fuzzy Logic Control

Python, Pytorch

 [code](#)

- Creating a new machine learning white-box model from scratch
- Creating a fuzzy inference system based on statistical inference
- Solving a system of integral equations using Pytorch

Machine learning of noise filtering of vibroacoustic linearly distributed sensor data


Python, TensorFlow

 [code](#)

- Creating an optimal signal filter for recognizing different types of activity

Galton board modeling

C++, Qt

 [code](#)

- Demonstration of the central limit theorem
- Creating a desktop application from scratch
- Embedding an engine for physical simulation

PROJECTS IN DEVELOPMENT

Adaptive metabolic model

Aug 2023 – Present

- Monte Carlo simulation
- Time series clustering

Adaptive control system with fuzzy logic based on Bayesian inference

Aug 2022 – Present

- The invention of the AI alignment method in RL
- Cross-entropy method for reinforcement learning
- Creating a greedy optimization algorithm for physical simulation
- Creating an analogue of the gradient descent algorithm in the function space