

Dmitry Nikitin

☎ +7(960)571-32-49 | @ dimnktn@gmail.com | in dimnktn | 🎧 Arist0craft | 📍 Moscow, Russia

SKILLS

Languages: C++, Python, Bash, SQL

Python Libraries: PyTorch, PySpark, FastApi, Pytest, Asyncio, CatBoost, XGBoost, LightGBM, Scikit-learn

DevOps and Workflow: Linux, Git, Hadoop, Spark, Docker, Apache Airflow, MLFlow, S3, CI/CD, Redis, Ansible

EXPERIENCE

Data Scientist

Dec 2023 – Present

Wildberries

One of top 3 marketplaces in Russia

- Designed the architecture of the Feature Store for ML models on ClickHouse, which accelerated the selection of data for training by 10%;
- Configured an automatic pipeline for predicting Lifetime Value (LTV) using Scikit-learn (linear regression, cross-validation), achieved accuracy of $R^2=0.81$, deployed the model via FastAPI and Streamlit, packing it into Docker containers;
- Developed an end-to-end ML pipeline for user clustering (RFM analysis, unsupervised learning methods), thereby identifying user clusters for marketing mailings and increasing revenue by 200+ thousand rubles;
- Implemented data drift detection using Apache Aiflow and Scipy, which allowed maintaining the quality of responses from 10+ models;
- Developed data quality tests (in Python + SQL) for an internal A/B testing platform, thereby identifying Z% of anomalies in experimental data, increasing the reliability of metrics;

Data Scientist

Mar 2023 – Dec 2023

VseInstrumenti.ru

Leading DIY e-commerce company in the CIS

- Built a number of ML and SARIMAX time series forecasting models with an average MAPE of 9% for key product metrics such as traffic analysis and B2C revenue;
- Conducted 30+ A/B tests for three product teams from MDE calculation to results summarization (using Stratification, Delta Method, Linearization, Bucketing and Bootstrap), which resulted in the implementation of 25+ new features and an increase in monthly revenue;
- Reduced data analysis time for ad-hoc queries for all product teams by 60-70% by building ETL architectures on Kubernetes, Apache Airflow and Google Data studio as a BI system;

Data Analyst

Mar 2021 – Mar 2023

AGIMA

The largest digital agency in Russia

- Built user churn prediction batch-model using CatBoost (ROC-AUC 0.82) and deployed via Airflow;
- Created a Tableau dashboard powered by Airflow pipelines, used by 6 product teams;
- Developed infrastructure for team of 4 analytics with GitLab CI/CD, JupyterHub and Metabase;

Researcher

Jun 2018 – Mar 2021

Joint Institute for Nuclear Research (JINR)

International nuclear physics research center

- Developed quantum mechanical two-body problem simulations using C++ and ROOT framework, improving numerical stability of particle interaction models by 15%;
- Implemented parallel computation algorithms with OpenMP, reducing simulation runtime for complex computation;

EDUCATION

Dubna State University

Sep 2016 – Jun 2020

Bachelor's degree in High Energy Physics

PROJECTS

Taxi Price Prediction Service | [*GitHub*](#)

- Developed a Dockerized taxi fare prediction service using CatBoost for return forecasting, deployed via FastAPI with a Telegram bot;

ML Algorithms implementation by hand | [*GitHub*](#)

- Implementation of classical ML algorithms using standart libraries with time/memory complexity and math description;