

Diveev Daniil



📍 Russia, Saint-Petersburg

✉️ daniildiveev@yandex.ru

👉 daniildiveev

Github

Hard Skills

AB-testing | Experimentation | Statistics

Hypothesis Testing | Machine Learning

Deep Learning | Data Visualisation

Data Manipulations | Causal Inference

Big Data | DB | ETL | DevOps

Programming: Python | Java | JavaScript

DB: PostgreSQL | MySQL | GreenPlum

Data Manipulations: Pandas | Polars
NumPy | Seaborn | Plotly

Data Visualisation: Tableau

BigData: PySpark | Hadoop

Machine Learning: Scikit-learn
Statsmodels

Deep Learning: PyTorch | TensorFlow
Keras | Transformers

DevOps: Docker | Kubernetes | Bash | Git

Education

Bachelor of Science

ITMO University

Neurotechnologies and programming

2022 - 2026

Saint-Petersburg, Russia

MTS Teta, data analytics school

2022

Moscow, Russia

Digital Banana

*Data Analysis and Machine Learning
course*

2019 - 2020

Saint-Petersburg, Russia

Languages

🇬🇧 English - C1

🇷🇺 Russian - Native

🇫🇷 French - A1

🇩🇪 German - B1

Work experience

DA at X5 Retail Group Targeting

since september 2023

- Designed over **100** ad hoc SQL queries, examined marketing campaigns logs, retrieved valuable information about segments behaviour and campaigns effectiveness
- Created **20+** marketing campaign reports giving information about crucial metrics change during campaign and treatment effectiveness using Python and Greenplum
- Achieved a **80% operational speed increase** of segment evaluation by creating a script using Python, PostgreSQL and KaitenAPI
- Increased **data extraction speed by 50%**, by developing a robust ETL pipeline using Python, Airflow and Greenplum

DA at X5 Retail Group CVM

june - september 2023

- Produced over **40** analytical reports for grocery suppliers (sales within region, average check, number of visits), utilizing PySpark for data processing and analysis
- Designed approximately **90 ad hoc queries**, examined customer behaviour during and after campaign, retrieved metrics (average check, conversion, turnover), created and presented reports using PySpark, Python and Excel
- Decreased campaign inference and goods buyers detection processes by **90%** by implementing automation scripts using PySpark and KaitenAPI
- Established approximately **20 target groups** for marketing campaigns, resulting in an average **revenue generation of 1 million rubles** each

DA at ITMO TRACK

january - march 2022

- Designed **85** ad hoc queries for insights and ideas about learning process effectiveness, visualised data, created reports using PostgreSQL and Python
- Conducted **2 customer developments**, totalling **~40 people** for ITMO.track product, giving the better understanding of target segment desires

Hackathons and researches

Hackathons

- Won Artificial Intelligence International Junior Contest 2021 (**multihop question answering using transformers**)
- Solved BCG Gamma hackathon case (**number of visits to Invitro clinics related to ticks prediction**) using ensemble of ARIMA models
- Solved Syntellic hackathon case (**prediction of molecule toxicity by its structure**) using gradient boostings fitted on GNN embeddings

Research

Title: "Electrochemical platform for detecting Escherichia coli bacteria using machine learning methods"

Description: Developed a machine learning algorithm to predict bacteria presence based on amperage values of different voltages