

ALEKSEI SHABANOV

Applied Data Scientist

📍 Berlin, Germany

✉ shabanoff.aleksei@gmail.com

🐙 @lexibender

🔗 alekseysh

DS with 6+ years of experience. Currently leading a team and organising the development of an open-source project.

WORK EXPERIENCE

Staff Data Scientist

NewYorker GmbH

📅 March 2021 – present

📍 Berlin, Germany

I led a team that works on vector representation for the **images** of the company's product catalog. I was responsible for research and development and also communicated with other teams and stakeholders to find the best way of using our model. Finally, we came up with exposing it as an internal API.

- Our colleagues boosted metrics in their data science projects via using our API.
- We also released the training pipeline as open-source library to improve our employer branding.

Applied Data Scientist

Neuromation Inc.

📅 March 2018 – October 2020

Working in a start-up allowed me to broaden my horizons and realize projects in machine learning consulting, for example:

- Built a **computer vision** system which was able to save doctors' time by checking if a box with surgical tools was correctly completed. The challenge was in solving the problem without retraining the model for the new tools (in a few-shot way).
- Created a **text** processing tool to help operators with extracting useful information from emails. It required to fulfill a project from communicating with the client and setting up the data labeling process to containerization and deploying the trained model.

Junior Data Analyst

3Red Trading, LLC

📅 June 2016 – March 2017

The company analyzed billions of transactions from the Chicago Mercantile Exchange. Based on statistical anomalies, I recognized traders who used automatic tools, which allowed our team to predict their behavior.

SKILLS AND ABILITIES

computer vision NLP tabular metric learning tracking

python PyTorch OpenCV pandas sklearn

Linux git CI/CD docker Kubernetes SQL

LANGUAGES

- English — Upper-Intermediate (B2)
- German — Elementary (A2)
- Russian — Native speaker

SIDE PROJECTS

Open-Source

OpenMetricLearning, author

OML is a new framework to train and validate models producing high-quality embeddings. It has a zoo of pretrained models, Config API, and a rich set of examples. I pay special attention to maintaining well-written documentation and covering the whole functionality with tests.

Catalyst, in top-10 contributors

It's a popular analogue of Keras for PyTorch (3k stars on GitHub, part of PyTorch ecosystem). Among other things, I've implemented a metric learning module for it, post.

Competitions

Challenge by Uber, 2nd out of 100+ teams

📅 11 Oct 2019 – 10 Feb 2020

This was a tabular competition with the following task: "*Predict when and where road incidents will occur next in Cape Town*". I organised the team and our strong sides were feature engineering for spatial data and a neat validation pipeline.

Kaggle competitions expert

JetBrains seminars at CS Center

Lecture about Person Re-Id problem

Video [ru].

EDUCATION

MSc in Applied Math and Physics

Saint-Petersburg State University

📅 2016 – 2019

Thesis title: "Semi-automatic cataloguing of images in a social network using machine learning methods", link.

BSc in Physics

Saint-Petersburg State University

📅 2012 – 2016

Thesis title "Statistical methods for the separating of interfering seismic waves".