

Artur Ilichev

✉@Lightingale ✉ilichev.art@gmail.com 🐙github.com/Lightingale 🔗linkedin.com/in/Lightingale

WORK EXPERIENCE

Yandex.Market (one of the largest marketplaces in Russia)

Data Scientist

October 2020 - Present

SQL Python MapReduce Catboost Toloka (crowdsourcing) A/B testing Mentorship

- Experimented with targets and features to improve the quality of user2item and item2item recommendation models (e.g., increased the number of orders of recommended similar products by **21%** using image and text embeddings).
- Implemented the automatic regular workflow for labeling search query redirects to product categories using crowdworkers (e.g., "OnePlus Buds" → "Headphones"). That improved the redirects recall by **14%** on popular search queries and by **8%** on all queries, with precision improved by **4%**.
- Developed an algorithm for finding trending goods to help users discover new items on the Market's main page.
- Analyzed the results of A/B tests and helped managers with making product decisions.
- Mentored my colleagues and participated in designing ML solutions for various projects.

Huawei Noah's Ark Lab (AI research center for Huawei Technologies)

Research Scientist (Natural Language Processing)

November 2019 - October 2020 (part-time)

NLP Python PyTorch TensorFlow

- Worked on the Transformer-based models distillation, studied the effect of distillation from multiple teacher models.
- Developed methods to improve the TinyBERT score on small datasets, e.g., on CoLA (linguistic correctness) by **15%**, using an improved augmentation process and an attention-based strategy for choosing the layer mapping function.
- Improved the distilled BERT robustness to random factors on the GLUE benchmark using multiple teachers.

Yandex (leading Russian search engine)

Machine Learning Engineer Intern

July 2019 - October 2019

Python NLP SQL Toloka (crowdsourcing) gensim Django ORM

- Developed a model to predict whether the dialogue is over (from collecting a dataset to writing production code).

Warden Machinery (Computer Vision Startup, Skolkovo Resident)

Machine Learning Engineer Intern

March 2019 - May 2019

Computer Vision Python PyTorch

- Developed heuristic methods for detecting illegal smoking based on the pose tracking and the Fourier transform.
- Trained and applied a neural network (MobileNetV2) for real-time special clothes classification on CCTV video.

Tinkoff Bank (the largest Russian digital bank)

Data Analyst Intern

July 2018

Data Visualization Python Plotly

- Performed visual data analysis, developed a model for prioritizing support queries to increase service productivity.

EDUCATION

Moscow Institute of Physics and Technology (MIPT)

Master's degree, Department of Applied Mathematics and Computer Science

GPA 4.8 / 5

Major in Applied AI (courses by Russian Venture Company)

Moscow Institute of Physics and Technology (MIPT)

Bachelor's degree, Department of Applied Mathematics and Computer Science

GPA 5 / 5

Major in Data Science (courses by Yandex School of Data Analysis)

ACHIEVEMENTS

Publications

- *RANLP 2021* - Multiple Teacher Distillation for Robust and Greener Models (A. **Ilichev**, N. Sorokin, V. Malykh, I. Piontkovskaya). [Link](#)

Kaggle (a platform for ML challenges) - Competitions Expert

- *2019* - Gendered Pronoun Resolution (Google Research) - 33rd out of 838 (**top 4%**, Silver medal).
- *2018* - Santander Value Prediction Challenge (Banco Santander) - 167th out of 4463 (**top 4%**, Silver medal).