

Pokonechnyy Eduard

GitHub LinkedIn

Email: pokonechnyy.ep@phystech.edu

Tel: +7-916-605-7888

Telegram: @celidos

Education

2021 – 2023, Master's Degree at MIPT

Department of Data Analysis (Yandex).

2019 – 2021, Yandex School of Data Analysis, Data Science track.

2015 – 2020, Bachelor's Degree at MIPT

Department of Image Recognition and Text Processing (ABBYY).

2013 – 2015, Advanced Education Scientific Center, Moscow State University

Work experience

◦ **Quality Visual Inspection System for a Brewery, Data Monsters,** *July 2019 – Present*

Problem: Build a system to automatically inspect cans running on a high-speed line and detect anomalies: dents, wrinkles and other anomalies, to prevent product loss and reputational risks.

Solution: Hardware set-up in the brewery + computer vision technologies to provide the image and can recognition process. My contribution is solving a variety of tasks with NN, including instance segmentation, anomaly/outlier unsupervised detection, image classification, image denoising, nearest neighbor search, speeding up computations and pipeline.

Used: Python, PyTorch, OpenCV, CUDA, Docker, k8s, helm

◦ **Determining the limits of applicability of the Neural Tangent Kernels (NTK) - Research project, iPavlov** *Feb 2022 – Jun 2022*

Problem: Evolution of an NN during training can be described by a Neural Tangent Kernel (NTK) [Jacot, 2018]. The task is to understand and try to expand the limits of applicability of the method to real problems, such as reducing the effective size of the training dataset, the problem of image classification.

Solution: Experimenting on different NN architectures (shallow nets, simple convolutional networks like Myrtle, Resnet50), hyperparameter tuning. Small paper preprint on review now (co-authored).

Used: Python, numpy, JAX, neural_tangents

◦ **Text chats intent classification and clusterization, Tinkoff bank** *Aug 2018 – Feb 2019*

Problem: Users come to bank sites with specific questions that they ask in the tech support chat and operators answer these questions. The task is to identify the basic structure of the requests received by the operators and to break these requests into groups to optimize the process with a lack of labeled data (clustering, classification).

Solution: Several state-of-the-art architectures (according to 2018) (fasttext, STC, ELMo etc). Strong data preprocessing.

Result: First-level topic clusterization is built, accepted by consumers with custom metrics.

Used: Python, numpy, gensim, scikit-learn, pandas

◦ **Junior Software Developer, RANEPa** *July 2018 – Dec 2018*

Successfully deployed an automatic service for monitoring the dynamics of prices for certain products in Russian food stores. C++, Python, Django, Web-scraping

◦ **Intern, Tester, ABBYY** *Summer 2017*

Deployed FlexiCapture app under Citrix, conducted experiments and performance testing.

Projects

- **iPavlov NLP course**

Workshops were based on materials from the Stanford course cs224n. Course covered both the basic knowledge of building neural networks and highly specialized techniques for NLP. The final team project: [NLP Question Answering system \(SQuAD\)](#). We used PyTorch to implement the model, based on attention mechanism and random dropout. [Poster](#) of our project.

- **Voice Application for 1C**

During [GlobalChangers](#) MIPT hackathon, our team developed small service for voice instrumentalization (Python, signal processing). Awarded with 1st place.

- **C++ computer graphics study project**

From scratch implemented with C++ a module for rendering images with realistic ray tracing and reflections. Processed lot of computational geometry.

- **Hackathons winner:**

- GlobalChangers 18' – news dump analysis, prediction of topic distribution
- MentorHack 18' – mutual mentoring and student search service
- Genesys 18' – project on multi-speaker speech separation

Skills & Technologies

- Python (numpy, scipy, pandas, tf/torch/keras, Web), Django, Flask, PyQt
- C++ (std11, stl, cmake), Qt, OpenGL
- Advanced algorithms and Data Structures
- Linux, Docker, SQL, Hadoop

- Probability theory, Statistics, Stochastic processes, Optimization Methods, Bayesian Methods in Machine Learning
- Advanced machine learning, Natural Language Processing, Theoretical Deep Learning
- Russian (native), English (intermediate)