

KONSTANTIN CHERNYSHEV

NLP Researcher and Engineer

@kdchernyshev+job@gmail.com

📍 Berlin, Germany

in kdchernyshev

🐙 k4black



EXPERIENCE

Machine Learning Researcher

📅 Jun 2024
– Present

🔗 [Toloka AI, Germany](#)

- Collected and published benchmark for text+visual university-level math (U-MATH, ACL 2025 accepted);
- Developed a substantial part of Agentic Platform for Human-AI collaboration, improving the quality on 30%+;

PyTorch HuggingFace GenAi Data Quality

NLP Data Scientist

📅 Jan 2022
– Sep 2022

🔗 [EPAM Systems, Serbia](#)

- Trained an unsupervised model for support ticket trends over time;
- Implemented a processing team predictive model, improving ticket processing efficiency by ~10s seconds per ticket for 100k tick./year.

Python FastText SkLearn PyTorch HuggingFace PyTest Docker Jenkins

Python Course Head, Lecturer

📅 Jan 2021
– Present

🔗 [Yandex School of Data Analysis, Russia](#)

- Led an Advanced Python course for 300-500 students yearly;
- Managed a team of 9 people, planning, review, tech supervision;
- Conducted lectures on 'Packages and Modules', 'Rust-bindings';
- Boosted assignment-checking system (e.g 6x acceleration).

Management Python Teaching Docker PyTest Gitlab-CI Github-Actions

Research Engineer, Project Manager

📅 Apr 2020
– Jan 2022

🔗 [Gorky AI \(startup\), Russia](#)

- Managed 2-4 developers per project, liaising with clients;
- Introduced Align flow, CI/CD, code reviews, improved teamwork;
- Developed a high-loaded 'roads quality' model on sensors data;
- Designed and developed a solution for scanned document analysis, trained a high mAP (~0.94) CV model for tables, imgs, and stamps.

Python PyTorch HuggingFace SkLearn PyTest ONNX Docker Gitlab-CI

Internships: 🔗 JetBrains Research Summer 2023, 🔗 LATNA Lab at Higher School of Economics 2019 - 2020;

EDUCATION

Master of Science (NLP)

📅 Sep 2022
– Aug 2024

🔗 [Groningen & Saarland Unis., Netherlands & Germany](#)

Erasmus Mundus "Language & Communication Technologies";
GPA: 8.7/10 +Assistant at Language Technology Project;
Thesis on Mechanistic Interpretability for LLM pruning.

Post Graduate 2-year Program (Data Science)

📅 Sep 2019
– Jul 2021

🔗 [Yandex School of Data Analysis, Russia](#)

Top Russian Data Science program with 3% acceptance rate;
GPA: 8.9/10; +Assistant at Python Course.

Bachelor of Science (Applied Math and CS)

📅 Sep 2017
– Jul 2021

🔗 [Higher School of Economics, Russia](#)

SUMMARY

Passionate Researcher with 6+ years of experience, doing XAI, Pruning and Agents;

- 6+ years of programming, 5+ of NLP Research experience in Startups, EPAM, JetBrains and Toloka AI;
- MsA with Honors at Erasmus Mundus LCT; BaSc with Honors in HSE, Russia;
- Completed 10+ ML research projects.

SKILLS

Data Science

NLP Agents XAI Pruning
Compression Speech Processing CV
RL

Frameworks

HuggingFace PyTorch SkLearn
Tensorflow NLPCore TorchText

Programming Languages

Python Rust C++ CUDA GoLang

Development Tools

git Docker Neptune.ai GitLab-CI
Github-Actions Ansible ONNX

Languages

Russian English German Serbian

PUBLICATIONS

- 🔗 LCT-1 at SemEval-2023 Task 10: Pre-training and Multi-task Learning for Sexism Detection and Classification [accepted, ACL-2023 workshop];
- 🔗 U-MATH: A University-Level Benchmark for Evaluating Mathematical Skills in LLMs [accepted, ACL-2025 workshop];

ACHIEVEMENTS

- Erasmus Mundus Scholarship 2022-2024; Honours Master's degree in LCT;
- Honours Bachelor's degree in CS;
- Largely improved Python course at YSDA, Top-1 by students' rating;
- Finished YSDA - Master's-level Data Science program, 3% acceptance rate.

PERSONAL

GPA: 9.1/10; +Assistant at Algorithms and Data Structures.
Thesis on Abstractive Text Summarization Knowledge Graphs.

Cheerful and sociable person, keen on
slackline and juggling, love music making.

Music Production

Juggling

Slackline