

Kuznetsov Andrey

INTERESTS

NLP / LLM / DL / CV / ML / TimeSeries / computer&math linguistics / Data Science / Digital Humanities

EDUCATION

ITMO / AI systems in arts and humanities

2022 – 2026

DLSchool MIPT / Deep Learning School

2023 – 2024

Summer School / Text data: automatic processing and analysis

July 2023 – August 2023

Sirius Educational Center / Linguistics

October 2021 – November 2021

JOB

NSS ITMO Lab / (Natural System Simulation)

assistant – winter 2023 – present time

Development of an open-source [framework](#) for automating the creation of CNN models that would set the parameters of layers and convolutions depending on the user. The product is used to predict weather conditions using n-dimensional time series.

T-Bank / (VoiceKit)

intern – March 2024 – July 2024

I worked in the TTS (text-to-speech) team on the task of automating the normalization of text and audio data as part of preprocessing before training. Development of the appropriate API for crowdsourcing and interacting with the NLP pipeline.

ACHIEVEMENTS

Selectel scholarship holder (January 2024 – July 2024)

The winner of the CODE-AI 2024 grant (July 2024 – November 2024)

Together with a team from the NSS laboratory, we won the AI CODE grant with the theme “*Library for automating the generation of convolutional neural networks for predictive modeling of natural environments.*”



SKILLS

Python: numpy / pandas / multiprocessing / asyncio / beautifulsoup / selenium / matplotlib / seaborn / sklearn / skimage / statsmodels / catboost / pytorch / natasha / spacy / nltk

SQL: PostgreSQL / pg4admin / MongoDB

R: stylo

Docker, Linux, Bash

LANGUAGES

English (Advanced)

FCE [certificate](#) (B2) 2021 год

CONTACTS

St. Petersburg, Russia
just.andrew.kd@gmail.com

[Telegram](#)
[GitHub](#)



НЕКОТОРЫЕ ПРОЕКТЫ

the BBBB team

We participate with our team in NLP hackathons at the national and international levels: <https://github.com/itmo-BBBB>

school-website-classification

Comparison of classification algorithms for school websites with different text vectorization, clusterization with an unknown number of clusters, statistical analysis, visualization in the web on streamlit

Promoting open source in ML:

- Participated in the stylometry project of the publication "System Block"
- Was a speaker at an Open Source [meetup](#)
- Parsed and uploaded my own datasets on [kaggle](#)
- Have a personal ML [channel](#)
- Write articles on [Habr](#).

ABOUT ME

Passionate about linguistics, I was a prize-winner of school Olympiads. Also, I was in Sirius at the Linguistics track, as well as at the NLP Summer School.

I am engaged in self-education. A regular student of the *MIPT Deep Learning School* courses, the *HSE Master's degree program "Machine Learning and High-load Systems"* and open online lectures of the SHAD

Stepik