

Кузнецов Андрей

СФЕРА ИНТЕРЕСОВ

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

ОБРАЗОВАНИЕ

ИТМО / Интеллектуальные системы в гуманитарной сфере

2022 – 2026 гг.

DLSchool МФТИ / Школа глубокого обучения

2023 – 2024 гг.

Летняя Школа / Текстовые данные: автоматическая обработка и анализ

Июль 2023 – Август 2023

Образовательный Центр Сириус / Лингвистика

Октябрь 2021 – Ноябрь 2021

ОПЫТ РАБОТЫ

Лаборатория NSS ИТМО / (Natural System Simulation)

лаборант – зима 2023 – настоящее время

Разработка open-source [библиотеки](#) для автоматизации создания CNN-архитектур, которые бы в зависимости от пользователя задавали параметры слоев и сверток. Продукт используется для прогнозирования метеословий с помощью n-мерных временных рядов.

T-Bank / (VoiceKit)

стажер – март 2024 – июль 2024

Работал в команде TTS (text-to-speech) над задачей автоматизации нормализации текстовых и аудио данных в рамках препроцессинга перед обучением. Писал соответствующий API для краудсорсинга и взаимодействовал с NLP пайплайном.

ДОСТИЖЕНИЯ

Стипендиат Selectel (январь 2024 – июль 2024)

Победитель гранта КОД-ИИ 2024 (июль 2024 – ноябрь 2024)

С командой из NSS лаборатории победили в гранте КОД-ИИ с темой “Библиотека для автоматизации генерации сверточных нейронных сетей для задач предсказательного моделирования природных сред”.



НАВЫКИ

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

ЯЗЫКИ

English (Advanced)

FCE [certificate](#) (B2) 2021 год

КОНТАКТЫ

г. Санкт-Петербург, Россия
just.andrew.kd@gmail.com

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



НЕКОТОРЫЕ ПРОЕКТЫ

команда BBBB

Со своей командой участвуем в NLP-хакатонах всероссийского и международного уровня: <https://github.com/itmo-BBBB>

school-website-classification

Сравнение классических алгоритмов классификации школьных сайтов с разной векторизацией текстов, кластеризация с неизвестным числом кластеров, статистический анализ, визуализация в web на streamlit

Продвигаю open-source в ML:

- Участвовал в проекте стилометрии издания "Системный блок"
- Был спикером на Open Source [митапе](#)
- Парсил и выкладывал собственные датасеты на [kaggle](#)
- Веду личный [канал](#) по ML
- Пишу статьи на [Хабр](#).

О СЕБЕ

Увлекаюсь лингвистикой, в школьные годы призер перечневых олимпиад. Был в Сириусе на треке ["Лингвистика"](#), а также в Летней Школе на [NLP направлении](#)

Занимаюсь самообразованием. Слушатель курсов *DLS МФТИ*, магистратуры ВШЭ *"Машинное обучение и высоконагруженные системы"* и открытых онлайн лекций *ШАД*

Stepik