



# ДАРЬЯ КАЛМЫКОВА

21 год г. Москва 1 год опыта

## ОПЫТ РАБОТЫ

- 2023-наст.вр., АДВ Лаборатория, Junior Data Scientist
  - EDA, обработка и подготовка данных для моделирования, построение эконометрических моделей, оптимизация и расширение функционала существующих моделей
- 2022-2023, кафедра МКИИТ “МТУСИ”, отдел аналитики
  - NLP : Кластеризация и последующая классификация элементов текстовой базы данных на классических алгоритмах машинного обучения, реализация NER.
  - Audio : классификация аудиозаписей по гендерному признаку и на наличие шума с использованием глубокого обучения и архитектуры Transformer.

## ДОПОЛНИТЕЛЬНОЕ ОБРАЗОВАНИЕ

- Stepik, курс “Введение в Data Science и машинное обучение”
- Stepik, курс “Нейронные сети и компьютерное зрение”
- Семен Козлов, курс “Deep Learning”
- Hugging Face, курс “NLP Course”
- Stepik, курс “Основы статистики”

## ЗНАНИЕ ЯЗЫКОВ

Русский	Родной язык
Английский	B2-Upper intermediate
Немецкий	Aufbau1-G3 - Aufbau1-G4

## ОБО МНЕ

Мне интересны разработка моделей МО и исследовательская деятельность в направлениях NLP, Audio, CV. Я бы хотела работать в команде, где приветствуется обмен знаниями и обучение у старших коллег.

Мои главные качества: целеустремленность, амбициозность и ответственность. Я хочу расширять свои профессиональные и учиться новому.



<https://github.com/Daryadare>



[ddarya.kalm@gmail.com](mailto:ddarya.kalm@gmail.com)



[@saddydare](https://twitter.com/saddydare)

СТЕК

★★★ Python

★★ C, C#

★ Java, C++

НАВЫКИ

Git

SQL

Docker

## ТЕХНОЛОГИИ

Word Embeddings, Bag of Words, TF-IDF, Ensemble modeling, PCA, K-Means, Transfer Learning, Transformers

## БИБЛИОТЕКИ

pandas, numpy, torch (torchvision, torchaudio), tensorflow, keras, CatBoost, scikit-learn, OpenCV, SciPy, nltk, word2vec, Gensim, matplotlib, seaborn, tqdm

## ОБРАЗОВАНИЕ

2020 – наст. время

“МТУСИ”, г. Москва, факультет Информационных Технологий, Бакалавр.

02.03.02 – Фундаментальная информатика и информационные технологии