

DENIS PISARENKO

Data Scientist

☎ 8985 082 33 17

✉ denisyuriveichpisarenko@gmail.com

🎂 July 11, 2000

📍 Moscow, Russia

🌐 [denisyuriveichpisarenko](https://denisyuriveichpisarenko.github.io)

📁 [ivaden](#)

SUMMARY

I am a data scientist with a strong background in linguistic research, experienced in data mining and building machine learning solutions for NLP-related tasks.

ACADEMIC RESOURCES

- [Academia.edu](#)
- [Google Scholar](#)

SKILLS

Languages: **Python** (pandas, numpy, matplotlib, seaborn, sklearn, pytorch, tensorflow, nltk, spacy, telebot)

R (tidyverse, mltools, caret, ggplot2)

SQL (filtering, aggregation of data, window functions, subqueries, derived tables, operations with strings)

Other: Latex, Git, Praat, ELAN, MS Excel, Kaggle, PythonAnyWhere

PROJECTS

PyTorch
BERT

Popular-Science Articles Difficulty Prediction

Prediction of difficulty of popular-science articles at N+1 website using multilingual BERT.

Sklearn
Spacy

Negation and Sentiment Analysis in BCS language

A study on the influence of different syntactic types of negation on model's performance in sentiment analysis task.

R
MLTools

Neg-Raising inference prediction

Testing different ML algorithms to predict an extremely rare linguistic inference.

PythonAnyWhere
Telebot

Nicknames Bot

A telegram bot created to give out the nicknames for students' grades table.

EDUCATION

2022 - curr.

HSE University, Moscow

Language Theory and Language Description

MA

2018 - 2022

HSE University, Moscow

Fundamental and Computational Linguistics

BA (with honor)

EXPERIENCE

2022 - 2023

Intern Researcher (Laboratory for Formal Models in Linguistics, HSE University)

- Data Collection and Statistical Analysis
- Research Results Presentation
- Experiment Design
- Field Research, Questionnaire Design

2022 - 2023

Tutor (HSE University)

- Teaching (Theory of Language, Introduction to Linguistics, Introduction to Syntax)
- Automatization of communication with students using telegram-bots

LANGUAGES

English - C1 (IELTS 7.5/9), **Russian** - native, **Korean** - A2, **Serbian** - B1, **German** - A2