

Olga Bystrova

NLP LEAD DATA SCIENTIST

Moscow, Russia / Worldwide

☎ (+7) 930-291-02-89 | ✉ bystrovaolgavl@gmail.com | 📷 [ovbystrova](#) | 🌐 [ovbystrova](#)

Work Experience

Sber

Moscow, Russia

SENIOR NLP DATA SCIENTIST

Aug. 2021 - Now

- **Deputy Team Lead** for 8 data scientists: currently processing team management, stakeholder communication, performance evaluation, candidate interviews
- **Led model development** efforts to build auto payment system based on contract recognition and payment documents evaluation. System processes > 5000 documents monthly
- Together with heads of related teams **introduced the process of project management** in short and long-term distance to balance the workload of the team and prevent delays on key projects
- **Mentored 3 team members**: supervised junior specialists in building and deploying pipeline with OCR/NER/Classification models; practiced regular code review and feedback loops
- **Collaborated with management and production teams** to create 4 demo stands of AI services for potential clients

Sber

Moscow, Russia

DATA SCIENTIST

Aug. 2020 - Aug. 2021

- Build and deployed **an ensemble of NER models** for contract recognition system to process 80% of documents automatically
- Worked on comprehensive **reviews system** from internal clients (sentiment analysis, topic modeling, custom classification models) to improve >90% business processes with AI
- Improved process of documents's report generation with **summarization model** as a first step saving from 5-10 minutes per document

Visiology

Moscow, Russia

JUNIOR DATA SCIENTIST

Sep. 2019 - Jun. 2020

- Virtual assistant ViTalk team: improved assistant's functionality to work with relative dates written by user
- Worked on speech2text algorithm that allows user to control BI dashboards by voice

Skills

Scientific Areas	DL, NLP (NER, Summarization, QA, Text Generation, Style Transfer), Audio (ASR, Speaker Diarization)
Deep Learning and Data Analysis	PyTorch, Sklearn, Numpy, Pandas, Transformers, NLTK, Spacy, Vosk, Nemo
Data Science workflow	Data Augmentation, Wandb, Docker, Flask, Streamlit
Languages	Russian (native), English (advanced), French (basic), German (basic)

Projects

InstructionNER: Generative NER

[github](#)

[HTTPS://GITHUB.COM/OVBYSTROVA/INSTRUCTIONNER](https://github.com/OVBYSTROVA/INSTRUCTIONNER)

July, 2022

- Unofficial implementation of InstructionNER paper
- Tuned T5 model to extract entities from input text

Neural Fake News Detector and text GANs

[HSE](#)

[HTTPS://GITHUB.COM/OVBYSTROVA/DPL](https://github.com/OVBYSTROVA/DPL)

June, 2020

- Used GPT-2 model to generate short texts with different sampling types
- Built several (BiLSTM with different embeddings, BERT models) binary classifiers to detect fake texts
- Compared accuracy of classifiers based on sampling type
- Trained language models as GANs with explored classifiers

Education

National Research University Higher School of Economics

Moscow, Russia

M.S. IN COMPUTATIONAL LINGUISTICS / NLP

Sep. 2020 - June. 2022

- GPA: 8.14
- Bachelor Thesis: "Neural Fake News Detector as a New Evaluation Metric for Text GANs".
- Master Thesis: "NLP Augmentation Techniques for Russian Language".
- Assisted on "Introduction to Data Analysis", "Introduction to Deep Learning" courses, Faculty of Computer Science

Certifications

2022 **Deep Learning Specialization**, deeplearning.ai
2020 **Deep Learning in NLP**, DeepPavlov, MPTI
2020 **Data Science Math Skills**, Duke University
2020 **SQL Interactive Trainer**, Far Eastern Federal University

Coursera
Moscow, Russia
Coursera
Stepic

Publications

Unreasonable Effectiveness of Rule-Based Heuristics in Solving Russian SuperGLUE Tasks

Dialogue conference, Moscow

COWRITER

June 2021

- <https://arxiv.org/abs/2105.01192>
- Implemented baseline solutions for RSG tasks: majority, random choice and random balanced choice
- Created heuristics for RUSSE, PARus, RuCoS tasks