

Daryna Dementieva

Education

Formal Education

- Nov 2019– Feb 2022 **PhD of Data Science**, *Skolkovo Institute of Science and Technology*, Computational and Data Science and Engineering Department.
Topic: Harmful information on social networks detection and combating
Adviser: Prof. Alexander Panchenko
- 2017–2019 **Master of Data Science**, *Moscow Institute of Physics and Technology*, Faculty of Applied Math and Physics, Department of Data Science, *GPA 4.7 out of 5.0*.
Thesis: News aggregation and personalization for professional groups
Adviser: Prof. Konstantin Vorontsov

Academic Mobility

- 2022 **Advanced Language Processing Winter School (ALPS) 2022**, *Grenoble, France*.
- 2021 **Advanced Language Processing Winter School (ALPS) 2021**, *Grenoble, France*.
- 2020 **The Machine Learning Summer School (MLSS) 2020**, *Tübingen, Germany*.
Poster Presentation: *Fake News Detection using Multilingual Evidence*

Publications (Google Scholar)

- 2022 **Varvara Logacheva***, **Daryna Dementieva***, **Sergey Ustyantsev**, **Daniil Moskovskiy**, **David Dale**, **Irina Krotova**, **Nikita Semenov**, **Alexander Panchenko**, *ParaDetox: Detoxification with Parallel Data*, Accepted to 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022).
- 2021 **David Dale**, **Anton Voronov**, **Daryna Dementieva**, **Varvara Logacheva**, **Olga Kozlova**, **Nikita Semenov** and **Alexander Panchenko**, *Text Detoxification using Large Pre-trained Neural Models*, The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021).
- 2021 **Daryna Dementieva**, **Sergey Ustyantsev**, **David Dale**, **Olga Kozlova**, **Nikita Semenov**, **Alexander Panchenko**, and **Varvara Logacheva**, *Crowdsourcing of Parallel Corpora: the Case of Style Transfer for Detoxification*, Proceedings of the 2nd Crowd Science Workshop: Trust, Ethics, and Excellence in Crowdsourced Data Management at Scale co-located with 47th International Conference on Very Large Data Bases (VLDB 2021).
- 2021 **Daryna Dementieva**, **Daniil Moskovskiy**, **Varvara Logacheva**, **David Dale**, **Olga Kozlova**, **Nikita Semenov**, and **Alexander Panchenko**, *Methods for detoxification of texts for the Russian language*, Multimodal Technologies and Interaction, 5(9) (MDPI 2021).

- 2021 **Daryna Dementieva, Alexander Panchenko**, *Cross-lingual evidence improves monolingual fake news detection.*, In the Proceedings of the ACL Student Research Workshop. Association for Computational Linguistics (ACL SRW 2021).
- 2021 **Daryna Dementieva, Daniil Moskovskiy, Varvara Logacheva, David Dale, Olga Kozlova, Nikita Semenov, and Alexander Panchenko**, *Methods for detoxification of texts for the Russian language*, In Proceedings of the 27th International Conference on Computational Linguistics and Intellectual Technologies (Dialogue 2021).
- 2020 **Daryna Dementieva, Alexander Panchenko**, *Fake News Detection using Multilingual Evidence*, IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA 2020).
- 2020 **Daryna Dementieva, Igor Markov, Alexander Panchenko**, *SkoltechNLP at SemEval-2020 Task 11: Exploring Unsupervised Text Augmentation for Propaganda Detection*, International Workshop on Semantic Evaluation (SemEval 2020).

Experience

- Mar 2021 - **NLP Researcher**, *Skoltech NLP Lab*, Moscow, Russia.
now Joint research with my PhD studies connected with machine translation (text style transfer).
- Jun–Sep 2018 **Data Science Intern**, *Beiersdorf*, Hamburg, Germany.
Working on models for automatic trends extraction from texts from social networks, e.g. Instagram, Medium.
- 2017–2019 **Data Scientist**, *Visiology*, Moscow, Russia.
- **Natural Language Assistant for BI Platform:**
Designing modules for dialogue system including creating new approaches for Russian language's specific tasks.
 - **Prompter for the operator on Sibur Neftekhim:**
Building machine learning models for analysis of production data, e.g. predicting time series with neural networks.

Additional Activities

- May 2019 **Speaker on DataFest 6**, Moscow, Russia, Gave a speech on the biggest Russian Data Science conference about NLP and virtual assistant for BI Platform.
- Oct 2018 - **Head of seminar NLP research group**, Moscow and Innopolis, Russia, Organizer
Aug 2019 and lecturer, author of the telegram channel @towards_nlp.

Hackathons & Business Case Championships

- November 2021 **Junction 2021**, *Helsinki, Finland*, SuperCell Uniting Empathy and Gaming Challenge, *3rd prize*: presenting research work of detoxification technology for real-world application..
- November 2019 **Junction 2019**, *Helsinki, Finland*, Ericsson Eco City competition, *2nd prize*: development of the taxi agent to minimize exhaust and maximize income..
- May 2018 **P&G IT Business Challenge**, *Athens, Greece*, One of three representatives of Russia at the international championship.

Apr 2017 **Management of the Future**, *Saint-Petersburg, Russia*, 1st prize: development of a prototype of the system for contactless payment by the user's fingerprint for VTB Bank .

Technical skills

Programming Languages

Experienced Python
Comfortable C, C++, Java

Tools

ML libraries sklearn, pytorch, Keras, Catboost
Analysis pandas, numpy, scipy
NLP NLTK, UDPipe, Faiss, fasttext, huggingface, BigARTM
Visualization matplotlib, plotly, bokeh, Grafana
Data bases PostgreSQL
VC Git

Languages

Russian Native
Ukrainian Native
English Advanced
German Intermediate