

Alexander Shirnin

alexamailwork@gmail.com • t.me/icecreamflavors

🌐 linkedin.com/in/alexander-shirnin • 🔄 25icecreamflavors • Kaggle profile

Education

National Research University HSE

MS in Business Analytics and Big Data Systems

Sep 2022 – July 2024

GPA: 7.8/10

National Research University HSE

BS in Applied Mathematics and Informatics

Sep 2018 – July 2022

GPA: 8.2/10

Work experience

HSE, Laboratory for Models and Methods of Computational Pragmatics

Oct 2021 - now

Intern researcher

- Conducting research on various NLP topics under the supervision of Dr. Ekaterina Artemova and Dr. Dmitry Ignatov.

HSE University

Jan 2023 - Mar 2023

Teaching

- Creating tasks for the data analysis exam.
- Teaching students Python language.

Publications

- A. Shirnin, N. Andreev, S. Potapova and E. Artemova, "Analyzing the Robustness of Vision & Language Models," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 32, pp. 2751-2763, 2024.
- A. Shirnin, N. Andreev, V. Mikhailov, E. Artemova. "Alpom at SemEval-2024 Task 8: Detecting AI-produced Outputs in M4." SemEval 2024: The 18th International Workshop on Semantic Evaluation, 2024.

Competitions

SemEval-2024 Task 8: Black-Box Machine-Generated Text Detection

2024

Workshop competition

- Developed a solution to detect boundaries in mixed text, where the first part is human-written and the second part is machine-generated. Used a LoRA-tuned LLM and an encoder model sequentially, creating an ensemble in an unconventional way.
- Finished 2nd out of 33 participants on the leaderboard.

Kaggle - LLM Science Exam

2023

Kaggle competition

- Trained DeBERTa models to answer difficult science-based questions (multi-choice test) written by a Large Language Model (GPT-3.5).
- Won a bronze medal on a Kaggle. Finished 228th out of 2664 teams, top 9%.

Toloka Visual Question Answering Challenge (WSDM Cup 2023)

2022

Conference competition

- Developed the solution for finding objects boxes in photos referred to by questions. Trained a VQA model for a visual grounding task and used ViT CLIP for blending predictions. Improved quality from the 0.21 IoU (organizers baseline) to the 0.73 IoU.
- Finished 5th out of 49 participants on the leaderboard.

CONDA competition for AJCAI 2022 workshop

2022

Workshop competition

- Achieved 2nd place on the Toxic Language Detection leaderboard, from the core level B conference (AJCAI). Trained JointBERT models ensemble for the toxic language detection (NER task).

Feedback Prize - Predicting Effective Arguments

2022

Kaggle competition

- Trained a DeBERTa models ensemble for classifying essays arguments in 3 categories.
- Won a bronze medal on a Kaggle. Finished 135th out of 1557 teams, top 9%.

Skills

Python libraries: PyTorch, HuggingFace, Pandas, Numpy, Scipy, Sklearn, XGBoost, Matplotlib.

Tools: Jupyter, L^AT_EX.

Kaggle: Competitions expert (highest ranking was 488/200,000 in global rating). 4 silver medals, 6 bronze medals.

Languages: English (advanced, IELTS 7.0), Russian (native speaker).