

# Elisei Stakovskii

NLP Engineer, Computational Linguist, Data Scientist

## Work Experience

---

Machine Learning Engineer (NLP), АО Эврика, October 2022 – Present:

- solving classification problem using topic modeling (BigARTM)
- data markup, NER model training
- writing a pipeline for filling the unified knowledge model (UKM) using NER models
- data collection, model training for the Text Classification task
- pipeline implementation for Text classification using NER models and Clasterization using custom Siemese Encoders
- data collection, training of MLM models for further fine-tuning for target tasks
- data collection, model training for the task of Relation Extraction for filling UKM
- mentoring and team leading of junior team members

Computational Linguist, Reverso, September 2020 – October 2022:

- expanding the list of German verbs for the Reverso Conjugator service using Wiktionary parsing: 80 verbs added
- a Python tool was written for text data preprocessing, cleaning and processing using the NeuroSpell API: a clean German corpus was prepared (30 thousand sentences) for further fine-tuning of the neural network model
- Python tool was written to collect up-to-date text data from websites for Spanish, French and German: 30,000 sentences collected for each language
- collecting and labeling data to improve the Ginger service: 20,000 sentences labeled
- creating and evaluating tests for NMT (Neural Machine Translation)
- localization and QA of Reverso and Ginger products

## Education

---

B.A HSE University  
(Philology, Linguistics and Traslation Studies),  
June 2020;  
GPA: 9.72/10



## Contacts

---

eistakovskii@gmail.com

<https://www.linkedin.com/in/elisei-stakovskii-203b0419a>

<https://github.com/eistakovskii>

<https://www.kaggle.com/eliseistakovskii>

<https://huggingface.co/ElStakovskii>

## Languages

---

Russian (Native)

English (C1, Advanced)

French (B2)

German (B2)

## Certificates

---

GRE (December, 2019): 314/340

TOEFL (October, 2021): 112/120

# Courses

---

- Machine Learning by Stanford, Coursera
- Deep Learning Specialization by DeepLearning.AI, Coursera
- Natural Language Processing Specialization by DeepLearning.AI, Coursera
- Machine Learning Engineering for Production (MLOps) Specialization by DeepLearning.AI, Coursera (in progress)

# Skill

---

**Theory:** Statistics, Probability Theory, Linear Algebra, Calculus

**NLP (Natural Language Processing) tasks:** sentiment analysis (logistic regression, naïve Bayes), POS tagging (Viterbi Algorithm), auto-complete (N-gram language model), auto-correct (minimum edit distance, dynamic programming), translation (locality-sensitive hashing, approximate k-nearest neighbor), Word2Vec (CBOW, Skip-gram), NER (LSTMs with linear layers, transformers), question duplicates identification ('Siamese' LSTM models), Transformer Summarizer, Question Answering, Chatbot (Reformer), Relation Extraction, Coreference Resolution, Text Generation

**Python Libraries:** NLTK, SpaCy, NumPy, Natasha, PyMorph, Selenium, Pandas, Transformers, Scikit-learn, Pytorch, TraX, Matplotlib, OpenCV, FastAPI, Flask, Requests, BigArtm, fasttext, gensim;

**Infrastructure:** Git, GitLab, Docker, SQL (MySQL), YouTrack, Jira, Linux (Ubuntu, Astra)

**Machine Learning:** Linear Regression, Logistic Regression, Neural Networks, Support Vector Machines, PCA, Random Forests, K-means

**Deep Learning:** RNNs (LSTM, GRU), GANs, Transformers (BERT, GPT, T5, Reformer; Finetuning, Adapters, Intermediary Pre-Training)