Taisiya Glushkova

glushkovato@gmail.com
glushkovato.github.io
+351 912 260 574

Education

PhD student in Computer Science, Instituto Superior Técnico — Lisbon, Portugal
Computer Science and Engineering Department; supervisor: Professor André F. T. Martins — 2020 – present
Collaborative project MAIA (Multilingual AI Agent Assistants) in the Carnegie Mellon Portugal Program
partnership involving Instituto de Telecomunicações, INESC-ID, Carnegie Mellon University and Unbabel.
Research interests: uncertainty estimation, probabilistic deep learning, nlp evaluation and interpretability

MSc in Data Science, HSE University — Moscow, Russia
Computer Science Department; supervisor: Professor Ekaterina Artemova — 2018 – 2020

BSc in Applied Mathematics and Computer Science, HSE University — Moscow, Russia
Computer Science Department; supervisor: Professor Ekaterina Artemova — 2014 – 2018

Experience

Instituto Superior Técnico — Lisbon, Portugal
PhD Researcher — Sep 2020 – present
- Uncertainty estimation for machine translation evaluation.

HSE University — Moscow, Russia
Research Internship (Part-time) — Oct 2019 – Aug 2020
- Question-answering for Russian language.

Unbabel — Lisbon, Portugal
Applied AI Research Summer Internship — Jul 2019 – Oct 2019
- Language model-based data filtering and selection for machine translation.

Digital Decisions — Moscow, Russia
Data Scientist (Python) — Feb 2019 – Mar 2020
- Hierarchical classification and routing of the customers messages.
- Named entity recognition for a speech assistant.

Sberbank — Moscow, Russia
Data Science Internship (Python) — Apr 2018 – Jun 2018
- Time series analysis and prediction. Forecasting demand for cash at ATMs.

Diginetica — Moscow, Russia
Junior Data Scientist/Engineer (Python) — Jun 2017 – Dec 2017
- Spell checker tool for search queries aimed to suggest corrections for misspelled words and offer synonyms.
- Snippets from product reviews: extraction of ultra-concise summaries of opinions over products.

Yandex — Moscow, Russia
Python Software Engineering Internship — Summer 2015
- Analysis of the personalities in the flow of search queries.
- Marking and classification of the various object definitions, ranking them in order of relevance for users.

Publications

*equal contribution

Uncertainty-Aware Machine Translation Evaluation — [ACL Anthology, slides]
Taisiya Glushkova, Chrysoula Zerva, Ricardo Rei, André F. T. Martins
Presented at the *Sixth Conference on Machine Translation (WMT21) at EMNLP 2021*
EMNLP Findings 2021

IST-Unbabel 2021 Submission for the Quality Estimation Shared Task — [wmt21]
Chrysoula Zerva, Daan van Stigt, Ricardo Rei, Ana C. Farinha, Pedro G. Ramos, José G. C. de Souza,
Taisiya Glushkova, Miguel Vera, Fabio Kepler, André F. T. Martins
EMNLP WMT 2021

Are References Really Needed? Unbabel-IST 2021 Submission for the Metrics Shared Task [wmt21]
Ricardo Rei, Ana C. Farinha, Chrysoula Zerva, Daan van Stigt, Craig Stewart, Pedro G. Ramos,
Taisiya Glushkova, André F. T. Martins, Alon Lavie
EMNLP WMT 2021

DaNetQA: a yes/no Question Answering Dataset for the Russian Language [arXiv]
Taisiya Glushkova*, Alexey Machnev*, Alena Fenogenova, Tatiana Shavrina, Ekaterina Artemova, Dmitry I. Ignatov
AIST 2020

Char-RNN and Active Learning for Hashtag Segmentation [arXiv]
Taisiya Glushkova, Ekaterina Artemova
CICLing 2019, published in Springer LNCS

Skills

Programming languages:	**Python**, R, C++, HTML, CSS, LATEX
Tools:	Jupyter Notebook, Google Colab, git, bash
Main libraries:	numpy, pandas, matplotlib, scipy, scikit-learn, nltk, gensim, fasttext, keras, pytorch, allennlp, deeppavlov, transformers
Languages:	English(Advanced), Russian(Native), German(A2), Portuguese(A1)

Teaching

Instituto Superior Técnico — Lisbon, Portugal
Teaching Assistant for Deep Learning course — Nov 2021 – present

HSE University — Moscow, Russia
Teaching Assistant for NLP course — Sep 2018 – Dec 2019
Teaching Assistant for Python course — Sep 2016 – Sep 2018

University Projects

Graduation Thesis (MSc), DaNetQA: a dataset for binary question answering in Russian — 2019 – 2020

Coursework, Detecting Uncertain Information in Natural Language Texts — 2018 – 2019

Graduation Thesis (BSc), Character Level Models for Hashtag Segmentation — 2017 – 2018

Coursework, Collecting Dictionaries for Sentiment Analysis of Reviews — 2016 – 2017

Coursework, Extracting Timestamps from News in English Language — 2015 – 2016

Real Time Tracker, Desktop App for managing tasks and projects in real time (analogue of Trello) — 2015 – 2016
Front end: HTML, CSS, jQuery, AngularJS framework
Back end: Python, Django, SQLite

Redis-server, Redis is an open source, in-memory data structure store, used as database, cache and message broker. — 2015
My goal was to implement the analogue by using C++ and boost library and to practise writing Unit-tests.

Extracurricular Activities

Participated in LxMLS (8th Lisbon Machine Learning School) — Summer 2018
Freshmen student advisor: helped students to choose courseloads and adjust to university life — 2015 – 2017
Organized university-wide events such as the Annual Open House — 2015
Hobbies: watercolor, hiking