Taisiya Glushkova

Home page / Google Scholar / Semantic Scholar / Github / Twitter / LinkedIn glushkovato@gmail.com

I'm a PhD student at SARDINE lab at Instituto Superior Técnico, supervised by André F. T. Martins and Chrysoula Zerva. I am broadly interested in Natural Language Generation Evaluation, Interpretability and Uncertainty Estimation. Currently, my focus is on developing methods and models that can quantify and communicate the uncertainty associated with decisions made by NLG models, so that users and stakeholders of these systems can make more informed decisions.

Education

PhD student in Computer Science, Instituto Superior Técnico

Lisbon, Portugal 2020 - present

Computer Science and Engineering Department; Supervisor: André F. T. Martins

Thesis title: Uncertainty Estimation and Robustness in Machine Translation Evaluation

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.

Collaborative project NextGenAl (Center for Responsible Al).

Research interests: uncertainty estimation, interpretability and robustness of machine translation evaluation

MSc in Data Science, HSE University

Moscow, Russia

Computer Science Department; Supervisor: Ekaterina Artemova

2018 - 2020

BSc in Applied Mathematics and Computer Science, HSE University

Moscow, Russia

Computer Science Department; Supervisor: Ekaterina Artemova

2014 - 2018

Experience

Carnegie Mellon University

Pittsburgh, USA

Research Internship

Jun 2022 - Aug 2022

Robust and interpretable machine translation evaluation. Supervisor: Alon Lavie.

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

Lisbon, Portugal

Jul 2019 - Oct 2019

• Question-answering for Russian language.

Unbabel

Applied Al Research Summer Internship

• Language model-based data filtering and selection for machine translation.

Digital Decisions Moscow, Russia Data Scientist Feb 2019 - Mar 2020

• Hierarchical classification and routing of the customers messages.

Named entity recognition for a speech assistant.

Diginetica Moscow, Russia

Junior Data Scientist/Engineer

Jun 2017 - Dec 2017

• Spell checker tool for search queries aimed to suggest synonymous corrections for misspelled words.

• 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 multiple object definitions, ranking them in order of relevance for users.

Selected Publications

BLEU Meets COMET: Combining Lexical and Neural Metrics Towards Robust Machine Translation Evaluation Taisiya Glushkova, Chrysoula Zerva, André F. T. Martins [ACL Anthology, poster] **EAMT 2023**

| Disentangling Uncertainty in Machine Translation Evaluation Chrysoula Zerva, Taisiya Glushkova , Ricardo Rei, André F. T. Martins EMNLP 2022 | | [ACL Anthology, poster] |
|--|--|--|
| Uncertainty-Aware Machine Translation Evaluation 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 | | [ACL Anthology, slides] |
| | bmission for the Metrics Shared Task rte Alves, Chrysoula Zerva, Ana C. Farinha, Coheur, André F. T. Martins | [ACL Anthology, poster] |
| Ricardo Rei, Marcos Treviso, Nuno M. | omission for the Quality Estimation Shared Task Guerreiro, Chrysoula Zerva, Ana C. Farinha, Christine Maroti, va, Alon Lavie, Luisa Coheur, André F. T. Martins | [ACL Anthology] |
| • | babel-IST 2021 Submission for the Metrics Shared Task a Zerva, Daan van Stigt, Craig Stewart, Pedro G. Ramos, ins, Alon Lavie | [ACL Anthology] |
| Char-RNN and Active Learning for Taisiya Glushkova, Ekaterina Artemo CICLing 2019, published in Springer Le | va | [arXiv] |
| Skills | | |
| Programming languages: Tools: Main libraries: Languages: | Python , R, C++, HTML, CSS, LATEX Jupyter Notebook, Google Colab, git, bash numpy, pandas, matplotlib, scipy, scikit-learn, keras, pytorch, tenglish(Advanced), Russian(Native), German(A2), Portuguese | |
| Teaching | | |
| Instituto Superior Técnico Teaching Assistant for Deep Learning course | | Lisbon, Portugal Nov 2021 – Feb 2022 |
| HSE University Teaching Assistant for NLP course Teaching Assistant for Python course | | Moscow, Russia Sep 2018 – Dec 2019 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 |
| Coursework, Extracting Timestamps to | from News in English Language | 2015 - 2010 |
| | managing tasks and projects in real time (analogue of Trello) | 2015 – 2016 2015 – 2016 |
| | managing tasks and projects in real time (analogue of Trello) | |