Taisiya Glushkova

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

I am a ML/NLP engineer, broadly interested in **LLMs, Evaluation, Interpretability and Uncertainty Estimation**. I recently completed a PhD, where my doctoral research focused on enhancing the robustness of Machine Translation evaluation models and developing methods that can quantify and communicate the uncertainty associated with their outputs. This experience has refined my skills in developing complex ML systems and ensuring their robustness and fairness.

Driven by the ongoing advancements in AI, I thrive on exploring the latest ML/DL tools, taking ownership of projects, and contributing to impactful work across various domains.

Based in Lisbon, Portugal. Hold a valid Portuguese residency permit, allowing work authorization. Open to relocation if needed.

Education

PhD in Computer Science and Engineering, Instituto Superior Técnico — Lisbon, Portugal
Thesis title: Uncertainty Estimation and Robustness in Machine Translation Evaluation — 2020 − 2024
Collaborative projects: MAIA (**Multilingual AI Agent Assistants**), NextGenAI (**Center for Responsible AI**)

MSc in Computer Science, HSE University — 2018 − 2020

BSc in Applied Mathematics and Computer Science, HSE University — 2014 − 2018

Work Experience

Instituto Superior Técnico, PhD Researcher — Lisbon, Portugal | 2020 − 2024
Led research on uncertainty estimation for efficient machine translation (MT) evaluation, resulting in the development of more robust evaluation models, under the supervision of André F. T. Martins.
- Contributed to extending COMET, state-of-the-art MT evaluation metric, with uncertainty measurement feature.
- Provided mentorship to junior PhD students, offering guidance on their research projects, enhancing their academic writing skills and facilitating regular feedback sessions.

Carnegie Mellon University, Research Intern — Pittsburgh, USA | Summer 2022
Developed robust and interpretable MT evaluation methods under the supervision of Alon Lavie.
- Collaborated with Unbabel on WMT Shared Tasks, securing second place in Metrics and winning the Quality Estimation task.
- Contributed to strategic planning and execution of research initiatives, aligning research goals with business objectives.

Digital Decisions, ML Engineer – Project Manager — Moscow, Russia | 2019 − 2020
Developed a hierarchical classification system for customer messages at Vkusvill, optimizing routing and response accuracy.
- Managed projects timelines, coordinated with stakeholders, and ensured deliverables met quality standards.

Unbabel, Applied AI Research Intern — Lisbon, Portugal | Summer 2019
Implemented language model-based data filtering for machine translation, improving data quality and translation accuracy.
- Enhanced translation model performance in several language pairs, leading to a higher customer satisfaction.

HSE University, Research Intern — Moscow, Russia | 2019 − 2020
Developed DaNetQA, a reproducible yes/no question-answering corpus, and applied transfer learning methods for task and language adaptation using multilingual fine-tuning and related sentence modeling tasks.

Diginetica, Junior Data Scientist/Engineer — Moscow, Russia | 2017
Created a tool for generating concise summaries from product reviews.
Developed a spell checker for search queries, enhancing correction suggestions for leading Russian eCommerce retailers.
- Improved search accuracy, enhancing user experience and engagement, resulting in 12% growth in search revenue.
- Conducted data-driven analysis and prioritized backlog tasks to ensure efficient management and delivery.

Yandex, Software Engineering Intern — Moscow, Russia | 2015
Implemented classification and ranking of multiple object definitions by user relevance.

Technical Skills

- Deep Learning, Machine Learning, Natural Language Processing.
- Fluent in Python, with minor experience with R and C++. Experience with DL frameworks: PyTorch, Keras.
- Proficient with common data science libraries such as NumPy, Pandas, Matplotlib, Scikit-learn, and HuggingFace.
- Comfortable with common data science tools, e.g. bash, git, Linux, slurm, Jupyter and SQL.

Selected Publications

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

Disentangling Uncertainty in Machine Translation Evaluation [ACL Anthology, poster]
Chrysoula Zerva, **Taisiya Glushkova**, Ricardo Rei, André F. T. Martins
EMNLP 2022

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

COMET-22: Unbabel-IST 2022 Submission for the Metrics Shared Task [ACL Anthology, poster]
Ricardo Rei, José G. C. de Souza, Duarte Alves, Chrysoula Zerva, Ana C. Farinha,
Taisiya Glushkova, Alon Lavie, Luisa Coheur, André F. T. Martins
EMNLP WMT 2022, Second place on the WMT 2022 Metrics shared task

CometKiwi: IST-Unbabel 2022 Submission for the Quality Estimation Shared Task [ACL Anthology]
Ricardo Rei, Marcos Treviso, Nuno M. Guerreiro, Chrysoula Zerva, Ana C. Farinha, Christine Maroti,
José G. C. de Souza, **Taisiya Glushkova**, Alon Lavie, Luisa Coheur, André F. T. Martins
EMNLP WMT 2022, Winner of the WMT 2022 QE shared task

Are References Really Needed? Unbabel-IST 2021 Submission for the Metrics Shared Task [ACL Anthology]
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

Teaching

Instituto Superior Técnico	Lisbon, Portugal
Teaching Assistant for Deep Learning course	2021 − 2022
HSE University	Moscow, Russia
Teaching Assistant for Natural Language Processing (NLP) course	2018 − 2019
Teaching Assistant for Python course	2016 − 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

Extracurricular Activities

Participated in a poster session at Center for Responsible AI	Demo Day	June 2024
Mentored junior PhD students, providing guidance on research projects and academic writing	2023-2024
Participated in a poster session at CMU Portugal 2022 Summit	November 2022
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