



LUDWIG-  
MAXIMILIANS-  
UNIVERSITÄT  
MÜNCHEN

CENTRUM FÜR INFORMATIONS- UND SPRACHVERARBEITUNG  
STUDIENGANG COMPUTERLINGUISTIK



## Thesis proposal

**Topic:** Are LLMs Robust to Typos Across Languages?

**Supervisor:** Yihong Liu

**Examiner:** Prof. Hinrich Schütze

**Level:** BSc/MSc

**Summary:** Humans are not immune to typos, or, typographical errors. Typos occur naturally and frequently in texts input by humans. In the context of NLP, an important research question to ask is whether LLMs are robust to typos. For example, the model should answer the two following queries similarly: (1) how is the weather and (2) how is the whether. Unfortunately, some LLMs are not robust even with one single character edit in English (Gan et al. 2024). In this study, we want to investigate the robustness of LLMs against typos across languages, not only Latin-script languages such as English. We aim to build an algorithm that can automatically generate a typo-version of an input text. The algorithm should account for multiple languages and the typos should be consistent with the common errors in the corresponding language. For example, in Chinese, randomly substituting a Chinese character to another one is a very rare occasion. Instead, it is a very common typo to substitute a character to another character with the same pronunciation, e.g., from 妈 (ma, which means mother) to 麻 (ma, which means horse) if people use Pinyin-based input method. With the algorithm, we are able to evaluate the LLMs' robustness against typos across languages and tasks.

**Requirements:** enthusiasm, good programming background (preferably python), good knowledge of NLP, a good command of PyTorch and HuggingFace.

### References:

- Esther Gan et al. (Nov. 2024). "Reasoning Robustness of LLMs to Adversarial Typographical Errors". In: *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. Ed. by Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen. Miami, Florida, USA: Association for Computational Linguistics, pp. 10449–10459. DOI: 10.18653/v1/2024.emnlp-main.584. URL: <https://aclanthology.org/2024.emnlp-main.584>