LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
CENTRUM FÜR INFORMATIONS- UND SPRACHVERARBEITUNG
STUDIENGANG COMPUTERLINGUISTIK

Thesis proposal

Topic: **Are Pixel-based Models More Robust Against Textual Perturbations?**

Supervisor: Yihong Liu

Examiner: Hinrich Schütze

Level: BSc / MSc

Summary: Unlike text-based language models that operate on subword tokenizations, pixel-based language models process text purely through its visual form. As a result, pixel-based models are expected to produce similar representations for texts that are visually similar, even when they differ at the character level. In contrast, text-based models may tokenize such inputs into entirely different subwords, yielding divergent representations. This project aims to systematically evaluate the robustness of pixel-based models to textual perturbations – such as typos, character swaps, and visually minimal edits – and to compare their behavior against standard text-based architectures.

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

References:

- Phillip Rust et al. (2023). *Language Modelling with Pixels*. arXiv: 2207.06991 [cs.CL]. URL: https://arxiv.org/abs/2207.06991
- Alberto Muñoz-Ortiz, Verena Blaschke, and Barbara Plank (2024). *Evaluating Pixel Language Models on Non-Standardized Languages*. arXiv: 2412.09084 [cs.CL]. URL: https://arxiv.org/abs/2412.09084