Evaluating the Performance of Anonymisation on Downstream Tasks

This master's thesis aims to investigate the impact of anonymisation techniques on the performance of downstream natural language processing (NLP) tasks. With the increasing concern for data privacy and regulations such as GDPR, organizations often employ anonymisation methods to protect sensitive information before sharing or analyzing data. However, the effectiveness of these techniques in preserving data utility while maintaining privacy is not well understood, especially in the context of NLP tasks. The central research question is: How does anonymisation affect the performance of NLP tasks, e.g., text classification, sentiment analysis, and named entity recognition?
To address this question, this research will involve:

1. Implementing and applying various anonymisation techniques to a diverse set of text datasets.
2. Conducting a systematic evaluation of the impact of anonymisation on the performance of different NLP tasks.
3. Analyzing the trade-offs between privacy and utility in anonymised data, considering factors like data quality, model accuracy, and robustness.
4. Identifying the most suitable anonymisation techniques for specific NLP applications based on the findings.

Expected outcomes of this research include a better understanding of the trade-offs between privacy and utility in the context of NLP, guidelines for selecting appropriate anonymisation techniques for different applications, and insights into potential improvements in privacy-preserving NLP.

Contact: **Fabian Schmidt** <fschm@kth.se>
Distributed Computing Group, SCS Division, KTH
Examiner: **Prof. Vladimir Vlassov** <vladv@kth.se>
Distributed Computing Group, SCS Division, KTH

Technical Requirements:

- Strong background in natural language processing and machine learning.
- Proficiency in programming languages like Python and experience with NLP libraries (e.g., NLTK, spaCy, TensorFlow, PyTorch).
- Familiarity with data anonymisation techniques, privacy-preserving methods, and ethical considerations.
- Statistical analysis and experimental design skills.
- Ability to work with large text datasets.

References:

- Mirshghallah, Fatemehsadat et al. “**Privacy in Deep Learning: A Survey.**” ArXiv abs/2004.12254 (2020).
- Iyadh Ben Cheikh Larbi et al. “**Clinical Text Anonymization, its Influence on Downstream NLP Tasks and the Risk of Re-Identification.**” ACL SRW. 2023.