Proceedings of the GermEval 2022 Workshop on Text Complexity Assessment of German Text

18th Conference on Natural Language Processing
KONVENS 2022

University of Potsdam
September 12, 2022

Edited by

Salar Mohtaj
Babak Naderi
Sebastian Möller

Preface

Text forms an integral part of exchanging information and interacting with the world. Along with the other types of content (e.g., image and video), textual content has been increased drastically in amount and importance during recent years. Text complexity (in the following used interchangeably with text readability) is one of the factors which affects a reader's understanding of text. A readability score is the mapping of a body of text to mathematical unit quantifying the degree of readability. We present the GermEval 2022 Workshop on Text Complexity Assessment of German Text. The task included developing Natural Language Processing (NLP) models to automatically assign a complexity score in the range of 1 to 7 to German texts. In other words, the shared task is a text regression task in which the output is continuous variables between 1 and 7. We received 84 submissions in the test phase from 10 teams. The results and the data sets can be found at the shared task website at https://qulab.github.io/text_complexity_challlenge/.

We are grateful to the *KONVENS* 2022 conference organizers for their support on the shared task. We would also like to show our gratitude to the participants of *GermEval* 2022 whose participation and effort made *GermEval* 2022 a great success. Finally we thank Kaspar Ensikat for his support for data preparation and Faraz Maschhur, Chuyang Wu, and Max Reinhard for developing the baseline model.

Potsdam, September 2022

The organizing committee

Organizers:
Salar Mohtaj (Technische Universität Berlin)
Babak Naderi (Technische Universität Berlin)
Sebastian Möller (Technische Universität Berlin)