

Evaluating the Simplification of Brazilian Legal Rulings in LLMs Using Readability Scores as a Target

Antônio Flávio Castro Torres de Paula, Celso Gonçalves Camilo

Institute of Informatics, Federal University of Goiás (UFG)

Abstract

This research focuses on evaluating the simplification of Brazilian legal rulings using large language models (LLMs). Legal documents are often complex, making it difficult for the general public to understand their content.

The study examines whether modern language models can simplify legal texts while preserving their original meaning, aiming to improve accessibility to legal information. The main question addressed by the study is whether large language models can automatically simplify Brazilian legal texts.

Our results with Llama-3 and Sabiá-2 show that while the complexity score decreases with higher readability targets, there is a trade-off with reduced semantic similarity.

Introduction

Legal documents, in their majority, have a complex language, characterized by the use of jargon and words that are infrequently used in common vocabulary, as well as domain-specific technical terms. These features hinder access to information for the Brazilian population and pose a challenge that must be addressed by the Brazilian justice system.

Most text simplification approaches require a ground truth, typically provided by human experts. However, the availability of resources and techniques for Brazilian Portuguese is limited, and even more so when considering the specific task of text simplification in the legal domain.

Objective

In this work, we evaluate the simplification of Brazilian legal rulings, using readability-controlled text modification task, and propose an evaluation approach that considers complex words specific to the evaluated domain.

Methodology

- ▶ Using readability-controlled text modification task
- ▶ FRES index as readability score adapted to Brazilian Portuguese
- ▶ Zero-shot prompting using FRES grade level to Brazil Education Levels
- ▶ Evaluation using WER and BERTScore to ensure lexical variability and maintain semantic meaning
- ▶ Evaluation approach using Complex Word Identification

Dataset

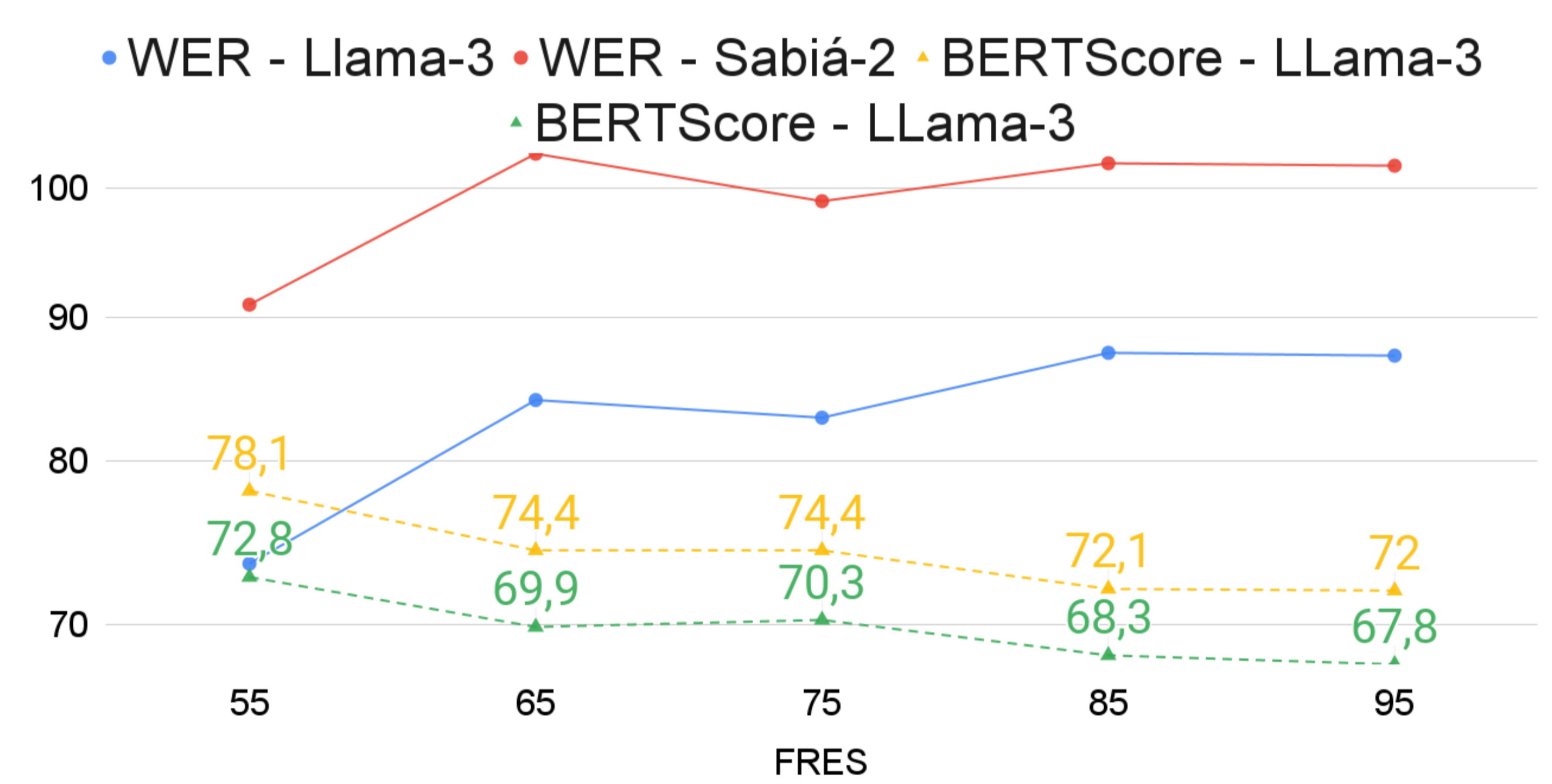
- ▶ 80k documents crawled from Court of Justice of Sao Paulo
 - ▷ Segmentation of Legal Rulings
 - ▷ 10k documents to evaluate text simplification
 - ▷ 195 unique judges
 - ▷ 3 year window time

Results

Individual-level metrics

Model	p(↑)	rmse(↓)	acc(↑)
Original data	0.0	44.35 ±14.63	2.24 ±6.52
Llama-3	61.05 ±37.38	24.46 ±10.13	10.89 ±15.32
Sabiá-2	22.76 ±49.64	20.41 ±7.29	16.51 ±15.14

Paraphrase metrics



Conclusion

This work applies the task of readability-controlled text modification, focusing on the simplification of legal texts.

We explore an approach based on complex word identification to evaluate the a text based on word complexity, indicating that the evaluated models have simplification capabilities and that there is a limit to this capacity, considering the proposed target scores.

In both evaluated models, Llama-3 and Sabiá-2, we observed that the complexity score decreases with higher readability scores, but with a reduction in the semantic similarity metric, highlighting the challenge of balancing simplification while preserving the main points of the original text.