Explainable Prediction of Text Complexity: The Missing Preliminaries for Text Simplification

Cristina Garbacea (garbacea@umich.edu)[1], Mengtian Guo[3], Samuel Carton[4] and Qiaozhu Mei[1,2]
[1] EECS Department, [2] School of Information, University of Michigan, Ann Arbor, MI, USA
[3] School of Information and Library Science, University of North Carolina, Chapel Hill, NC, USA
[4] CS Department, University of Colorado, Boulder, CO, USA

ACL 2021

Background & Motivation

Text simplification aims to reduce the language complexity of highly specialized textual content so that it is accessible for readers who lack adequate literacy skills, such as children, people with low education, people who have reading disorders or dyslexia, and non-native speakers.

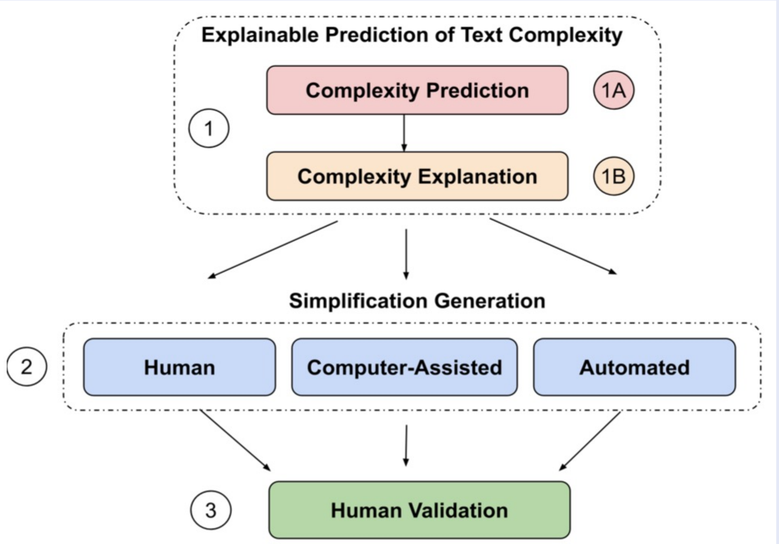
Mismatch between language complexity and literacy skills is a critical source of bias and inequality: 18 years of education are required on average to properly understand clinical trial descriptions (Wu et al, 2016).

Text simplification has considerable potential to improve the fairness and transparency of text information systems, for eg. in healthcare, education, for kids and English learners.

However, the definition of text simplification in the literature has never been transparent. End-to-end neural network models have been widely adopted to directly generate the simplified version of input text, usually functioning as a black-box.

Text Simplification Pipeline

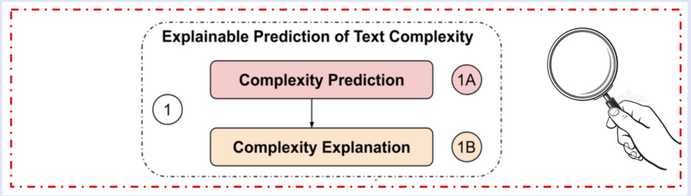
We show that the general problem of text simplification can be formally decomposed into a compact and transparent pipeline of modular tasks to ensure transparency and explainability of the process.



Explaining the rationale behind the simplification decisions may support generalization in unseen scenarios (Doshi-Velez and Kim, 2018).

Explainable Prediction of Text Complexity

It is critical to explain the rationale behind simplification decisions, especially behind a black-box model. No prior work has addressed the explainability of text complexity prediction:



1A) predict whether a given piece of text needs to be simplified, **1B)** if yes, identify complex parts of the text.
These two tasks can be solved *separately*, using lexical or deep learning methods, or *jointly* through an end-to-end, explainable predictor.

Experimental Setup & Results

Datasets: aligned complex-simple sentence pairs from three different domains - news, Wikipedia, scientific papers

Dataset	Training	Validation	Test
Newsela	94,208 pairs	1,129 pairs	1,077 pairs
WikiLarge	208,384 pairs	29,760 pairs	59,546 pairs
Biendata	29,700 pairs	4,242 pairs	8,486 pairs

Complexity Prediction: wide portfolio of deep and shallow binary classifiers to distinguish complex sentences from simple ones.

Classifier	Newsela	WikiLarge	Biendata
NB n-grams	73.10 %	62.70 %	84.30 %
NB enriched features	73.10 %	63.10 %	86.00 %
LR n-grams	75.30 %	71.90 %	89.60 %
LR enriched features	76.30 %	72.60 %	91.70 %
SVM n-grams	75.20 %	71.90 %	89.50 %
SVM enriched features	77.39 %	70.16 %	88.60 %
RF n-grams	71.50 %	71.50 %	84.60 %
RF enriched features	74.40 %	73.40 %	87.00 %
LSTM (word-level)	73.31 %	71.62 %	89.87 %
CNN (word-level)	70.71 %	69.27 %	89.05 %
CNN (char-level)	78.83%†	74.88 %	88.00 %
CNN (word & char-level)	75.90	74.00 %	92.30 %
Extractive Adversarial Networks	72.76 %	71.50 %	88.64 %
ULMFiT	**80.83%****	74.80 %	94.17 %
BERT	77.15 %	**81.45%****	94.43 %
XLNeT	78.83%†	73.49 %	**95.48%****

Complexity Explanation: how well the complex parts of a sentence are highlighted to explain complexity prediction

Explanatory Model	Complexity Explanation
LIME & LR	Their fatigue changes their voices , but they 're still on the freedom highway .
LIME & LSTM	Their fatigue changes their voices , but they 're still on the freedom highway .
SHAP & LR	Their fatigue changes their voices , but they 're still on the freedom highway .
Extractive Networks	Their fatigue changes their voices , but they 're still on the freedom highway .
Simple sentence	Still , they are fighting for their rights .

How to evaluate Complexity Explanation?

ideally, ground truth annotations at the token level are available; if not, as a proxy, all tokens w_i in complex sentence d which are absent in simple sentence d' are candidate words for deletion or substitution

Dataset	Explanation Model	P	R	F1	TER	ED 1.5
Newsela	Random	0.515	0.487	0.439	0.985	13.825
	AoA lexicon	**0.556**	0.550	0.520	0.867	12.899
	LR Features	0.522	0.250	0.321	0.871	12.103
	LIME & LR	0.535	0.285	0.343	0.924	12.459
	LIME & LSTM	0.543	**0.818**	**0.621**	0.852	11.991
	SHAP & LR	0.553	0.604	0.546	0.848	12.656
	Extractive Networks	0.530	0.567	0.518	**0.781**	**11.406**
WikiLarge	Random	0.412	0.439	0.341	1.546	17.028
	AoA lexicon	0.427	0.409	0.357	1.516	16.731
	LR Features	0.442	0.525	0.413	0.993	17.933
	LIME & LR	0.461	0.509	0.415	**0.988**	18.162
	LIME & LSTM	**0.880**	0.470	0.595	1.961	25.051
	SHAP & LR	0.842	**0.531**	**0.633**	1.693	22.811
	Extractive Networks	0.452	0.429	0.359	1.434	**16.407**
Biendata	Random	0.743	0.436	0.504	1.065	12.921
	AoA lexicon	0.763	0.383	0.475	1.064	13.247
	LR Features	0.796	0.257	0.374	0.979	10.851
	LIME & LR	**0.837**	0.466	0.577	0.982	**10.397**
	LIME & LSTM	0.828	0.657	0.713	**0.952**	16.568
	SHAP & LR	0.825	0.561	0.647	0.979	11.908
	Extractive Networks	0.784	**0.773**	**0.758**	0.972	10.678

Benefit of Complexity Prediction

A smart end-to-end simplification model should not further simplify input if already simple enough. However, current best pre-trained simplification models ACCESS (Martin, 2020) and DMLTL (Guo, 2018) incorrectly simplify >70% and >90% of out-of-sample simple sentences:

- In Ethiopia, HIV disclosure is low → In Ethiopia , HIV is low (ACCESS)
- Healthy diet linked to lower risk of chronic lung disease → Healthy diet linked to lung disease (DMLMTL)

Explainable prediction of text complexity avoids such pitfalls and yields better out-of-sample performance (30-70% error reduction)!

Dataset	Sentence Pairs	Metric	ACCESS	DMLMTL
Newsela	No complexity prediction (simplify everything)	ED	4.044	12.212
		TER	0.175	1.611
		FED	0.016	0.170
	With complexity prediction (predicted simple: no change)	ED	**2.631** (-35%)	**8.677** (-29%)
		TER	0.089 (-49%)	1.149 (-29%)
		FED	0.006 (-63%)	0.066 (-61%)
WikiLarge	No Complexity Prediction (simplify everything)	ED	5.857	16.920
		TER	0.208	2.328
		FED	0.004	0.143
	With Complexity Prediction (predicted simple: no change)	ED	**4.021** (-31%)	**10.566** (-38%)
		TER	0.132 (-37%)	1.452 (-38%)
		FED	0.002 (-50%)	0.049 (-66%)
Biendata	No Complexity Prediction (simplify everything)	ED	3.796	9.030
		TER	0.254	1.348
		FED	0.033	0.131
	With Complexity Prediction (predicted simple: no change)	ED	**1.887** (-50%)	**5.249** (-42%)
		TER	0.114 (-55%)	0.819 (-39%)
		FED	0.009 (-73%)	0.051 (-61%)

Key Takeaway

Major motivation of text simplification: improve fairness,transparency it is critical to explain the rationale behind the simplification decisions behind a black-box model to reduce the out-of-sample error

This work is in part supported by the National Science Foundation under grant numbers 1633370 and 1620319 and by the National Library of Medicine under grant number 2R01LM010681-05.