

cases	doc_1		doc_2		decision	id
	authors	<ul style="list-style-type: none"><li>Liu, Yang</li></ul>	authors	<ul style="list-style-type: none"><li>Liu, Yang Jr</li></ul>	DUPLICATES	294
	title	Assessing text readability and quality with language models	title	Assessing text readability and quality with language models		
	publication_date	2020-01-01 00:00:00	publication_date	2020-01-01 00:00:00		
	source	SupportedSources.CORE	source	SupportedSources.CORE		
	journal		journal			
	volume		volume			
	doi	None	doi	None		
	urls	<ul style="list-style-type: none"><li>https://core.ac.uk/download/288487248.pdf</li></ul>	urls	<ul style="list-style-type: none"><li>https://core.ac.uk/download/288487365.pdf</li></ul>		
	id	id6387010666205165583	id	id-4629993051214889769		
	abstract	Automatic readability assessment is considered as a challenging task in NLP due to its high degree of subjectivity. The majority prior work in assessing readability has focused on identifying the level of education necessary for comprehension without the consideration of text quality, i.e., how naturally the text flows from the perspective of a native speaker. Therefore, in this thesis, we aim to use language models, trained on well-written prose, to measure not only text readability in terms of comprehension but text quality. In this thesis, we developed two word-level metrics based on the concordance of article text with predictions made using language models to assess text readability and quality. We evaluate both metrics on a set of corpora used for readability assessment or automated essay scoring (AES) by measuring the correlation between scores assigned by our metrics and human raters. According to the experimental results, our metrics are strongly correlated with text quality, which achieve 0.4-0.6 correlations on 7 out of 9 datasets. We demonstrate that GPT-2 surpasses other language models, including the bigram model, LSTM, and bidirectional LSTM, on the task of estimating text quality in a zero-shot setting, and GPT-2 perplexity-based measure is a reasonable indicator for text quality evaluation	abstract	Automatic readability assessment is considered as a challenging task in NLP due to its high degree of subjectivity. The majority prior work in assessing readability has focused on identifying the level of education necessary for comprehension without the consideration of text quality, i.e., how naturally the text flows from the perspective of a native speaker. Therefore, in this thesis, we aim to use language models, trained on well-written prose, to measure not only text readability in terms of comprehension but text quality. In this thesis, we developed two word-level metrics based on the concordance of article text with predictions made using language models to assess text readability and quality. We evaluate both metrics on a set of corpora used for readability assessment or automated essay scoring (AES) by measuring the correlation between scores assigned by our metrics and human raters. According to the experimental results, our metrics are strongly correlated with text quality, which achieve 0.4-0.6 correlations on 7 out of 9 datasets. We demonstrate that GPT-2 surpasses other language models, including the bigram model, LSTM, and bidirectional LSTM, on the task of estimating text quality in a zero-shot setting, and GPT-2 perplexity-based measure is a reasonable indicator for text quality evaluation		
	versions		versions			