

Beautiful or handsome? Neural language models try their hand at word substitution

13 May 2021



Credit: CC0 Public Domain

Researchers from Skoltech and their colleagues have run a first of its kind large-scale computational study of the most advanced neural language models to see how they handle lexical substitution, a crucial task in natural language processing. The paper was presented at the 28th International Conference on Computational Linguistics (COLING-2020).

Lexical substitution is the [task](#) of replacing a particular word in a sentence with another word that is somehow related to the original word and still appropriate for the context. For instance, in a sentence "Who killed Laura Palmer?" the word killed can be replaced with a synonym murdered. In "I'm the king of the world!" we can replace king with a hypernym (a word with a broader meaning) ruler. And in "You're gonna need a bigger boat," sea enthusiasts can arguably replace boat with a meronym hull, using the word that means a part of something to refer to the whole object.

Lexical substitution is quite easy for native human speakers of a language, but it is a much harder task for machines that have to do [natural language](#)

processing (NLP). They might need it for word sense induction, that is, identifying the particular meaning of a word in context, correcting spelling based on the meaning of the word, and even more complex tasks such as paraphrasing or simplifying a text. For this purpose, language models based on [deep neural networks](#) are built that can produce a number of substitutes for a target word based on its immediate left and right context.

Assistant professor Alexander Panchenko of Skoltech and his colleagues from Samsung Research Center Russia, HSE University, and Lomonosov Moscow State University decided to run a competition of sorts between five neural language models. They tested the models on two tasks: lexical substitution itself and word sense induction (when a machine has to distinguish between a bank of a river and a bank as a financial institution).

The researchers believe their results may be useful for NLP practitioners. They were able to show, among other things, which models tend to produce semantic relations of which types (synonyms, hypernyms and so on mentioned earlier) and that additional information about the target word can boost the quality of lexical substitution quite substantially—or significantly, if you're looking for a synonym here.

"First of all, our results in lexical substitution may be useful for [language](#) learning (replacing words with their simpler equivalents). Second, it may be useful for augmentation of textual data for training [neural networks](#), as similar augmentation methods are common in computer vision but not so common in text analysis. Another obvious application is writing assistance—automatic suggestion of synonyms and text reformulation," Panchenko says.

More information: Nikolay Arefyev et al. Always Keep your Target in Mind: Studying Semantics and Improving Performance of Neural Lexical

Substitution, *Proceedings of the 28th International Conference on Computational Linguistics* (2021).
[DOI: 10.18653/v1/2020.coling-main.107](https://doi.org/10.18653/v1/2020.coling-main.107)

Provided by Skolkovo Institute of Science and Technology
APA citation: Beautiful or handsome? Neural language models try their hand at word substitution (2021, May 13) retrieved 24 May 2022 from <https://techxplore.com/news/2021-05-beautiful-handsome-neural-language-word.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.