Grammatical Error Correction

Overview

This technology deals with automatic correction of different kinds of errors in written text, including spelling, punctuation, grammatical, word choice, and stylistic errors.

An illustrating example:

The problems ~~bring some effect on~~ *affect engineering design* ~~from~~ *in two* ~~aspect~~ *aspects, independent innovation and engineering application.*

Background

There are more non-native speakers of English than there are native English speakers. However, learning English is difficult and many conventions of the English language are hard to acquire for non-native speakers.

Existing grammar checking software mostly relies on manually written rules and does not achieve high accuracy.

Approach

We use a neural approach by learning how to correct text from large collections of text examples. We take advantage of error-annotated learner data, such as the NUS Corpus of Learner English (NUCLE) that we collected, and very large non-learner texts (about 90 billion words) from the Web.

Achievements

When evaluated on a standard benchmark test data set (CoNLL-2014), our approach achieves the highest accuracy published to date (based on $F_{0.5}$ score computed by the official M2 scorer) [Chollampatt and Ng, AAAI 2018].

Applications

This technology can be used in automatic grammar checker and educational language learning software.