

Deep learning boosts Google Translate tool

Internet giant's latest service employs neural networks to cut error rate by 60% the company says.

Daide Castelvechi

27 September 2016



Thomas Samson/AFP/Getty Images

Google Translate — here shown on a mobile phone — will use deep learning to improve its translations between texts.

Google's online translation service, Google Translate, will soon be using a new algorithm that is entirely based on [deep learning](#), the company [announced](#) on 27 September. The algorithm, which is also described in a paper posted to the preprint server [arXiv¹](#), is the first widely-available computer system for translating languages that relies on the increasingly popular AI technique. Compared to the firm's existing service, the algorithm reduces errors by around 60%, Google computer scientists say.

A Chinese-to-English service that uses the algorithm is now being used on the Google Translate mobile and web-based apps, and Google says that it will roll out other languages over the next few months.

The advance is yet another example of the success of deep learning, which has helped crack major AI problems in recent years by combining artificial neural networks — layers of computational units that mimic the way neurons connect in the brain — with enormous data sets. Most notably, the technique has beaten other machine approaches to [image recognition](#) and [game-playing](#). Google has now applied the same approach to language translation, to create what it calls a Neural Machine Translation system (NMTS). "From the input to the output, it's entirely done by one neural network," says Quoc Le, a computer scientist at Google in Mountain View, California, who helped to develop the NMTS.

The NMTS borrows from others' advances in machine learning and adds a few methodological novelties, says Yoshua Bengio, a computer scientist at the University of Montreal in Canada who has read the arXiv paper. "On a first look, it appears that they have pulled most of the known tricks," he says. The team's main achievement is to show that "with solid engineering and a well-designed architecture, neural-machine translation can far outpace classical methods for machine translation", he says, with "pretty amazing results".

The algorithm "really improves the state-of-the-art in many ways", says Jürgen Schmidhuber, a computer scientist at the University of Lugano (also known as the USI) in Manno, Switzerland.

Machine translation

Until now, Google Translate's use of artificial neural networks has been limited, says Charina Choi, a company spokesperson. For the most part, its algorithms analysed text word by word, learning to associate corresponding words in different languages by scouring millions of existing translations, such as in documents from the United Nations or the European Parliament.

The NMTS also learns by analysing existing translations; as it does so, it tweaks connections between artificial neurons in a way that will improve its performance. But it analyses sentences by first breaking up each word into 'word segments', an idea that came from team member Mike Schuster, also at Google in Mountain View, who had employed it in speech-recognition software.

"Somehow, in some representation inside the neural network, the segments can combine to represent meaning," says Le. This could be similar to the way that neural networks perform visual tasks such as face recognition: they start from the individual pixels in an image and work up through increasingly complex features such as edges, geometric patterns and so on.

The same neural network that analysed the text then produces a translation. To improve speed, the company runs the system on computer chips that it designed specifically for machine learning. AlphaGo, the program that [beat a top human player at the game of Go](#) earlier this year, used similar hardware.

Performance verdict

To evaluate the translation system's performance, the Google researchers took sentences from Wikipedia and from news articles and put NMTS-made translations (between a handful of different language pairs) side by side with corresponding ones made by the company's old system and by human translators. The team then had human evaluators score the quality of the translations in a blind test.

Chinese-to-English translation, which is notoriously difficult, showed marked improvements, but still lagged compared to the algorithm's translations among Indo-European languages. For some other language pairs, the accuracy of the NMTS approached that of human translators, although the authors caution that the significance of the test was limited by its sample of well-crafted, simple sentences.

Schmidhuber thinks that machines will only be able to truly match or beat humans at translation once they are able to combine different sensory inputs. "Today, they only see sentences such as 'the cat fell from the tree'," Schmidhuber says. "In the future, they will also see videos of cats falling from trees, and they will control robots that can see, hear, move and manipulate objects, and feel pain through pain sensors, and relate their experiences to texts."

Nature | doi:10.1038/nature.2016.20696

References

1. Wu, Y. *et al.* Preprint at <http://arxiv.org/abs/1609.08144v1> (2016).