Reading comprehension: Alibaba model may get better marks than you

January 17 2018, by Nancy Owano

Credit: Marina Shemesh/public domain

Take some soothing blueberry juice. Or dust off your worry beads. Or anything else you do for calm when you read about artificial intelligence beating humans in mind games. Here comes another.

AI surpassed humans on a reading test. Not just any reading test. At center stage is The Stanford Question Answering Dataset (SQuAD).

Specifically, this is a reading comprehension dataset carrying questions that were posed by crowdworkers on Wikipedia articles.

Alibaba developed an artificial intelligence model that emerged victorious on this test, having scored better than humans in reading and comprehension. This was developed by Alibaba's Institute of Data Science of Technologies.

Xinhua carried Alibaba's statement that "This is the first time that a machine has outperformed humans on such a test."

How did they do it? According to Xinhua, "Alibaba explained that its AI was able to win because its neural network model is based on the Hierarchical Attention Network which enables the AI to read from 'paragraphs to sentences to words' in order to identify phases that can have potential answers."

Robert Fenner on Monday in Bloomberg said the test was "considered one of the world's most authoritative machine-reading gauges." Carl Engelking on Monday in *Discover* described it as "an arduous test" of a machine's natural language processing skills.

So, with this test, they are talking about over 100,000 question-answer pairs on over 500 articles.

The AI scored 82.44, just past the 82.304 that humans achieved.

The Alibaba model used natural-language processing, which, said Fenner, mimics human comprehension of words and sentences.

Engelking in *Discover* brought it to light. "'What changes the mineral content of a rock?' These questions are a level higher than simply scanning for basic facts, and they require algorithms to process a large

amount of information regarding context, sequences and relationships before providing an accurate answer."

Why this matters: For Engelking, "2018 marks the year that, by one measure, machines surpassed humans' reading comprehension abilities."

But wait.

Jamie Condliffe, *MIT Technology Review*, sought to remind people about something in "The Download" on Monday. Alibaba's AIs outperformed humans in the comprehension test, but, he added, tough natural language challenges are still facing machines.

"This isn't comprehension the way humans think of it," said Condliffe. "It's neat, but the AI doesn't really understand what it reads—it doesn't know what 'British rock group Coldplay' really is, besides it being the answer to the Super Bowl question. And there are far harder language problems that humans still beat computers at."

Meanwhile, Alibaba, known as a Chinese internet giant, joins others "in a race to develop AI that can enrich social media feeds, target ads and services or even aid in autonomous driving," wrote Fenner in Bloomberg.

In a statement, scientist Luo Si spelled out potential applications. "The technology underneath can be gradually applied to numerous applications such as customer service, museum tutorials and online responses to medical inquiries from patients, decreasing the need for human input in an unprecedented way."

Citation: Reading comprehension: Alibaba model may get better marks than you (2018, January