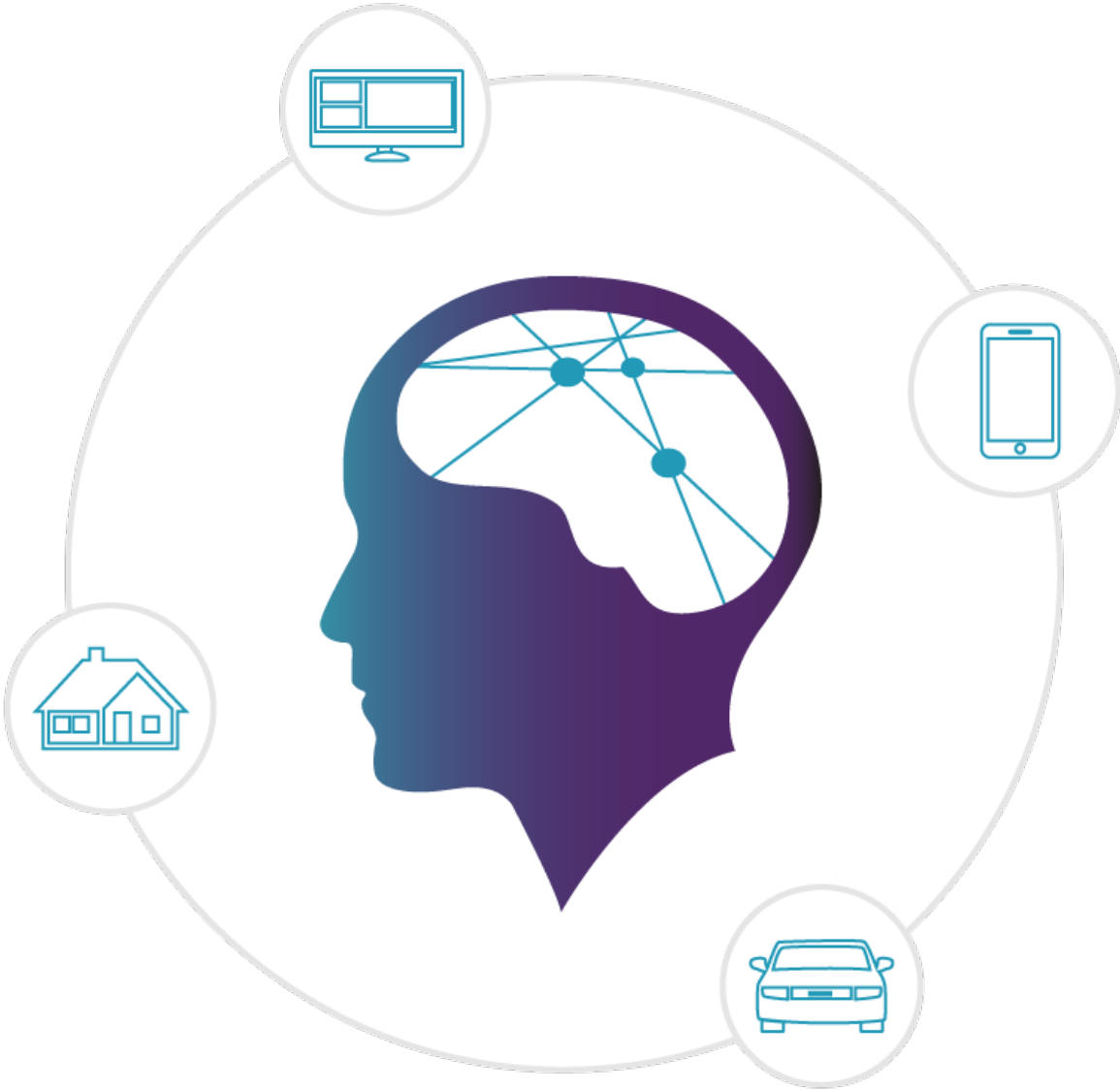


Maluuba researchers try algorithm on Harry Potter text

March 31 2016, by Nancy Owano



Credit: Maluuba

(Tech Xplore)—When have you heard this before: Machine learning advances the collective intelligence of humans... a lot has been written about machine learning vis-a-vis computer applications, the Internet of things, and in time people may become more and more familiar with a company called Maluuba. They have something they refer to as "[machine](#) intelligence technology."

The Ontario, Canada-based company is pushing their competence in advanced research in machine reading comprehension.

Maluuba certainly drew the keen attention of *MIT Technology Review's* senior editor for AI, Will Knight, earlier this week for a reason other than the company's enthusiasm. Knight said they are to something with a substantially practical application potential. The title of his article was "Software that reads Harry Potter might perform some wizardry."

What that means is that the Maluuba team is training deep-learning algorithms to answer questions about small amounts of [text](#). Knight would like us to think about that. He wrote, "what if a computer could read a few dozen pages of text, like the manual for a new microwave, and then answer [questions](#) about how it works? Sign me up."

Comprehension of text by machines, at a near-human level, has been generally a top research goal for [natural language processing](#). Maluuba recently demonstrated in a video just what they have achieved in a reading of Harry Potter, displaying their tech's ability to answer questions correctly. The company has taken up a challenge, in the form of a very difficult task for computers, which is that of comprehending.

The [company](#), reported *Cantech Letter* on Tuesday, "has applied an algorithm to the text of J.K. Rowling's bestselling novel *Harry Potter and the Philosopher's Stone*, along with several hundred other children's stories, to read text in such a way that it can then answer questions afterward."

Knight pointed out that "A fundamental challenge with language is that the words used to represent different concepts are arbitrary, so it is more difficult to draw connections between them than it is for images."

Knight said the team made progress "with an algorithm that can read text and answer questions about it with impressive accuracy." He said it could in time help computers to comprehend documents.

They said in a video that so far achieved state of the art results in MCTest but they decided to test their machine comprehension system on more challenging stories. (Microsoft defines MCTest as "a freely available set of 660 stories and associated questions intended for research on the machine comprehension of [text](#).")

In their paper, they referred to MCTest as "a complex but data-limited comprehension benchmark, whose multiple-choice questions require not only extraction but also inference and limited reasoning. Their paper noted numerous hand-engineered features which cannot be trained. Their model, significantly, "can be trained end-to-end with backpropagation."

They turned to something which everyone knows, they said, Harry Potter, linguistically, much more complex than the MCTest that their model was trained on. The vocabulary was different too. Nonetheless, they said, they did pretty well in answering questions about the text. This was a multiple choice format of options in response to [questions](#).

"A Parallel-Hierarchical Model for Machine Comprehension on Sparse Data" is their paper and it was submitted this month on arXiv. Authors are Adam Trischler, Zheng Ye, Xingdi Yuan, Jing He, Phillip Bachman and Kaheer Suleman.

The authors wrote, "We have presented the novel Parallel-Hierarchical model for machine comprehension, and evaluated it on the small but complex MCTest. Our model achieves state-of-the-art results, outperforming several feature-engineered and neural approaches. Working with our model has emphasized to us the following (not necessarily novel) concepts, which we record here to promote further empirical validation.

The authors said the training wheels approach—initializing neural networks to perform sensible Heuristics—appears helpful for small datasets. They also said that reasoning over language was challenging, but "easily simulated in some cases."

According to reports, the company is building a research facility in artificial intelligence. Their general [goal](#) lies in teaching machines to think, reason and communicate. The research lab will be led by Maluuba's CTO, Kaheer Suleman, and will be staffed by 13 deep learning research scientists, said *BetaKit*.

More information: www.maluuba.com/

© 2016 Tech Xplore

Citation: Maluuba researchers try algorithm on Harry Potter text (2016, March 31) retrieved 20 April 2024 from <https://techxplore.com/news/2016-03-maluuba-algorithm-harry-potter-text.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.