

Natural language processing software evaluates middle school science essays

October 11 2022, by Mariah Chuprinski



Credit: CC0 Public Domain

Students may soon have another teacher in the classroom, but from an unlikely source: artificial intelligence (AI). In two recent papers, computer scientists at Penn State vetted the effectiveness of a form of

AI known as natural language processing for assessing and providing feedback on students' science essays. They detailed their results in the publishing arm of the [International Society for the Learning Sciences Conference](#) (ISLS) and in the Proceedings of the [International Conference on Artificial Intelligence in Education](#) (AIED).

Natural language processing is a subfield of computer science where researchers convert the written or spoken word into computable data, according to principal investigator Rebecca Passonneau, Penn State professor of computer science and engineering.

Led by Passonneau, the researchers who worked on the ISLS paper extended the abilities of an existing [natural language](#) processing tool called [PyrEval](#) to assess ideas in [student](#) writing based on predetermined, computable rubrics. They named the new software PyrEval-CR.

"PyrEval-CR can provide [middle school students](#) immediate feedback on their science essays, which offloads much of the burden of assessment from the teacher, so that more writing assignments can be integrated into middle school science curricula," Passonneau said. "Simultaneously, the software generates a summary report on topics or ideas present in the essays from one or more classrooms, so teachers can quickly determine if [students](#) have genuinely understood a science lesson."

The beginnings of PyrEval-CR date back to 2004, when Passonneau worked with collaborators to develop the [Pyramid method](#), where researchers annotate source documents manually to reliably rank written ideas by their importance. Starting in 2012, Passonneau and her graduate students worked to automate Pyramid, which led to the creation of the fully automated PyrEval, the precursor of PyrEval-CR.

The researchers tested the functionality and reliability of PyrEval-CR on hundreds of real middle school science essays from public schools in

Wisconsin. Sadhana Puntambekar, professor of educational psychology at the University of Wisconsin-Madison and a collaborator on both papers, recruited the science teachers and developed the science curriculum. She also provided historical student essay data that was needed to develop PyrEval-CR before deploying it in classrooms.

"In PyrEval-CR, we created the same kind of model that PyrEval would create from a few passages by expert writers but extended it to align with whatever rubric makes sense for a particular essay prompt," Passonneau said. "We did a lot of experiments to fine-tune the software, then confirmed that the software's assessment correlated very highly with an assessment from a manual rubric developed and applied by Puntambekar's lab."

In the AIED paper, researchers relay the technical details on how they adapted the PyrEval software to create PyrEval-CR. According to Passonneau, most software is designed as a set of modules, or building blocks, each of which has a different function.

One of PyrEval's modules automatically creates the assessment model, called a pyramid, from four to five reference texts written to the same prompt as the student essays. In the new PyrEval-CR, the assessment model, or computable rubric, is created semi-automatically before students even receive an essay prompt.

"PyrEval-CR makes things easier for teachers in actual classrooms who use rubrics, but who usually don't have the resources to create their own rubric and test whether it can be used by different people and achieve the same assessment of student work," Passonneau said.

To evaluate essays, students' sentences must first be broken down into individual clauses and then converted to fixed-length sequences of numbers, known as vectors, according to Passonneau. To capture the

meaning of clauses in their conversion to vectors, an algorithm called weighted text matrix factorization is used. Passonneau said the algorithm captured the essential similarities of meaning better than other tested methods.

Researchers adapted another algorithm, known as weighted maximal independent set, to ensure PyrEval-CR selects the best analysis of a given sentence.

"There are many ways to break down a sentence, and each sentence may be a complex or a simple statement," Passonneau said. "Humans know if two sentences are similar by reading them. To simulate this human skill, we convert each rubric idea to vectors, and construct a graph where each node represents matches of a student vector to rubric vectors, so that the software can find the optimal interpretation of the student [essay](#)."

Eventually, researchers hope to deploy the assessment software in classrooms to make assigning and assessing science essays more practical for teachers.

"Through this research, we hope to scaffold student learning in science classes, to give them just enough support and feedback and then back off so they can learn and achieve on their own," Passonneau said. "The goal is to allow STEM teachers to easily implement writing assignments in their curricula."

In addition to Passonneau and Puntambekar, the other contributors to the ISLS paper are: Purushartha Singh and ChanMin Kim, Penn State School of Electrical Engineering and Computer Science; and Dana Gnesdilow, Samantha Baker, Xuesong Cang and William Goss, University of Wisconsin-Madison. In addition to Passonneau and Puntambekar, the other contributors to the AIED paper are Mohammad Wasih, Penn State School of Electrical Engineering and Computer

Science; Singh, Kim and Cang.

More information: Purushartha Singh et al, Automated Support to Scaffold Students' Written Explanations in Science, *Artificial Intelligence in Education* (2022). [DOI: 10.1007/978-3-031-11644-5_64](https://doi.org/10.1007/978-3-031-11644-5_64)

Provided by Pennsylvania State University

Citation: Natural language processing software evaluates middle school science essays (2022, October 11) retrieved 10 April 2024 from <https://techxplore.com/news/2022-10-natural-language-software-middle-school.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.