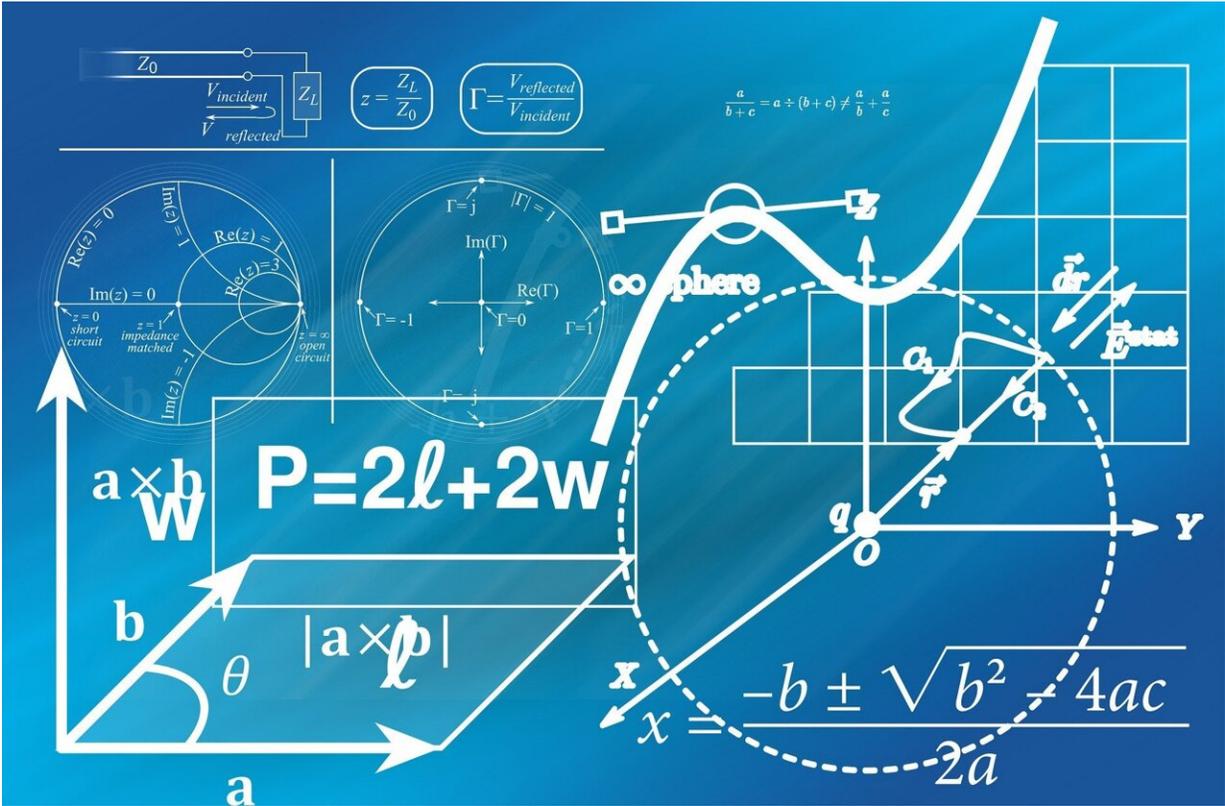


New algorithm aces university math course questions

August 3 2022, by Adam Zewe



Credit: Pixabay/CC0 Public Domain

Multivariable calculus, differential equations, linear algebra—topics that many MIT students can ace without breaking a sweat—have consistently stumped machine learning models. The best models have only been able

to answer elementary or high school-level math questions, and they don't always find the correct solutions.

Now, a multidisciplinary team of researchers from MIT and elsewhere, led by Iddo Drori, a lecturer in the MIT Department of Electrical Engineering and Computer Science (EECS), has used a [neural network model](#) to solve university-level math problems in a few seconds at a human level.

The model also automatically explains solutions and rapidly generates new problems in university math subjects. When the researchers showed these machine-generated questions to [university students](#), the students were unable to tell whether the questions were generated by an algorithm or a human.

This work could be used to streamline content generation for courses, which could be especially useful in large residential courses and [massive open online courses](#) (MOOCs) that have thousands of students. The system could also be used as an automated tutor that shows students the steps involved in solving undergraduate math problems.

"We think this will improve higher education," says Drori, the work's lead author who is also an adjunct associate professor in the Department of Computer Science at Columbia University, and who will join the faculty at Boston University this summer. "It will help students improve, and it will help teachers create new content, and it could help increase the level of difficulty in some courses. It also allows us to build a graph of questions and courses, which helps us understand the relationship between courses and their pre-requisites, not just by historically contemplating them, but based on data."

The work is a collaboration including students, researchers, and faculty at MIT, Columbia University, Harvard University, and the University of

Waterloo. The senior author is Gilbert Strang, a professor of mathematics at MIT. The research appears this week in the *Proceedings of the National Academy of Sciences*.

A "eureka" moment

Drori and his students and colleagues have been working on this project for nearly two years. They were finding that models pretrained using text only could not do better than 8 percent accuracy on high school math problems, and those using graph neural networks could ace machine learning course questions but would take a week to train.

Then Drori had what he describes as a "eureka" moment: He decided to try taking questions from undergraduate math courses offered by MIT and one from Columbia University that had never been seen before by a model, turning them into programming tasks, and applying techniques known as program synthesis and few-shot learning. Turning a question into a programming task could be as simple as rewriting the question "find the distance between two points" as "write a program that finds the difference between two points," or providing a few question-program pairs as examples.

Before feeding those programming tasks to a [neural network](#), however, the researchers added a new step that enabled it to vastly outperform their previous attempts.

In the past, they and others who've approached this problem have used a neural [network](#), such as GPT-3, that was pretrained on text only, meaning it was shown millions of examples of text to learn the patterns of natural language. This time, they used a neural network pretrained on text that was also "fine-tuned" on code. This network, called Codex, was produced by OpenAI. Fine-tuning is essentially another pretraining step that can improve the performance of a machine-learning model.

The pretrained model was shown millions of examples of code from online repositories. Because this model's training data included millions of natural language words as well as millions of lines of code, it learns the relationships between pieces of text and pieces of code.

Many math problems can be solved using a computational graph or tree, but it is difficult to turn a problem written in text into this type of representation, Drori explains. Because this model has learned the relationships between text and code, however, it can turn a text question into code, given just a few question-code examples, and then run the code to answer the problem.

"When you just ask a question in text, it is hard for a machine-learning model to come up with an answer, even though the answer may be in the text," he says. "This work fills in the that missing piece of using code and program synthesis."

This work is the first to solve undergraduate math problems and moves the needle from 8 percent accuracy to over 80 percent, Drori adds.

Adding context

Turning math questions into programming tasks is not always simple, Drori says. Some problems require researchers to add context so the neural network can process the question correctly. A student would pick up this context while taking the course, but a neural network doesn't have this background knowledge unless the researchers specify it.

For instance, they might need to clarify that the "network" in a question's text refers to "[neural networks](#)" rather than "communications networks." Or they might need to tell the model which programming package to use. They may also need to provide certain definitions; in a question about poker hands, they may need to tell the model that each deck contains 52

cards.

They automatically feed these programming tasks, with the included context and examples, to the pretrained and fine-tuned neural network, which outputs a program that usually produces the correct answer. It was correct for more than 80 percent of the questions.

The researchers also used their model to generate questions by giving the neural network a series of math problems on a topic and then asking it to create a new one.

"In some topics, it surprised us. For example, there were questions about quantum detection of horizontal and vertical lines, and it generated new questions about quantum detection of diagonal lines. So, it is not just generating new questions by replacing values and variables in the existing questions," Drori says.

Human-generated vs. machine-generated questions

The researchers tested the machine-generated questions by showing them to university students. The researchers gave students 10 questions from each undergraduate math course in a random order; five were created by humans and five were machine-generated.

Students were unable to tell whether the machine-generated questions were produced by an algorithm or a human, and they gave human-generated and machine-generated questions similar marks for level of difficulty and appropriateness for the course.

Drori is quick to point out that this work is not intended to replace human professors.

"Automation is now at 80 percent, but automation will never be 100

percent accurate. Every time you solve something, someone will come up with a harder question. But this work opens the field for people to start solving harder and harder questions with machine learning. We think it will have a great impact on [higher education](#)," he says.

The team is excited by the success of their approach, and have extended the work to handle math proofs, but there are some limitations they plan to tackle. Currently, the model isn't able to answer questions with a visual component and cannot solve problems that are computationally intractable due to computational complexity.

In addition to overcoming these hurdles, they are working to scale the [model](#) up to hundreds of courses. With those hundreds of courses, they will generate more data that can enhance automation and provide insights into course design and curricula.

More information: Iddo Drori et al, A neural network solves, explains, and generates university math problems by program synthesis and few-shot learning at human level, *Proceedings of the National Academy of Sciences* (2022). [DOI: 10.1073/pnas.2123433119](https://doi.org/10.1073/pnas.2123433119)

Provided by Massachusetts Institute of Technology

Citation: New algorithm aces university math course questions (2022, August 3) retrieved 18 April 2024 from <https://techxplore.com/news/2022-08-algorithm-aces-university-math.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.