MATHEMATICS

Machine learning comes up against unsolvable problem

Researchers run into a logical paradox discovered by famed mathematician Kurt Gödel.

BY DAVIDE CASTELVECCHI

A team of researchers has stumbled on a question that is mathematically unanswerable because it is linked to logical paradoxes discovered by Austrian mathematician Kurt Gödel in the 1930s.

The mathematicians, who were working on a machine-learning problem, show that the question of 'learnability' — whether an algorithm can extract a pattern from limited data — is linked to a paradox known as the continuum hypothesis. Gödel showed that this cannot be proved either true or false using standard mathematical language. The latest result were published on 7 January (S. Ben-David *et al. Nature Mach. Intel.* **1**, 44–48; 2019).

"For us, it was a surprise," says Amir Yehudayoff at the Technion–Israel Institute of Technology in Haifa, who is a co-author on the paper. He says that although there are a number of technical maths questions that are known to be similarly 'undecidable', he did not expect this phenomenon to show up in a seemingly simple problem in machine learning.

John Tucker, a computer scientist at Swansea University, UK, says that the paper is "a heavyweight result on the limits of our knowledge", with foundational implications for both mathematics and machine learning.

NOT ALL SETS ARE EQUAL

Researchers often define learnability in terms of whether an algorithm can generalize its knowledge. The algorithm is given the answer to a 'yes or no' question — such as "Does this image show a cat?" — for a limited number of objects, and then has to guess the answer for new objects.

Yehudayoff and his collaborators arrived at their result while investigating the connection between learnability and 'compression', which involves finding a way to summarize the salient features of a large set of data in a smaller set of data. The authors discovered that the information's ability to be compressed efficiently boils down to a question in the theory of sets — mathematical collections of objects such

Austrian mathematician Kurt Gödel.

as the sets in Venn diagrams. In particular, it relates to the different sizes of sets containing infinitely many objects.

Georg Cantor, the founder of set theory, demonstrated in the 1870s that not all infinite sets are created equal: in particular, the set of integer numbers is 'smaller' than the set of all real numbers, also known as the continuum. (The real numbers include the irrational numbers, as well as rationals and integers.) Cantor also suggested that there cannot be sets of intermediate size — that is, larger than the integers but smaller than the continuum. But he was not able to prove this continuum hypothesis, and nor were many mathematicians and logicians who followed him.

Their efforts were in vain. A 1940 result by Gödel (which was completed in the 1960s by US mathematician Paul Cohen) showed that the continuum hypothesis cannot be proved either true or false starting from the standard axioms — the statements taken to be true — of the theory of sets, which are commonly considered to be the foundation for all of mathematics.

Gödel and Cohen's work on the continuum hypothesis implies that there can exist parallel mathematical universes that are both compatible with standard mathematics — one in which the continuum hypothesis is added to the standard axioms and therefore declared to be true, and another in which it is declared false.

LEARNABILITY LIMBO

In the latest paper, Yehudayoff and his collaborators define learnability as the ability to make predictions about a large data set by sampling a small number of data points. The link with Cantor's problem is that there are infinitely many ways of choosing the smaller set, but the size of that infinity is unknown.

The authors go on to show that if the continuum hypothesis is true, a small sample is sufficient to make the extrapolation. But if it is false, no finite sample can ever be enough. In this way, they show that the problem of learnability is equivalent to the continuum hypothesis. Therefore, the learnability problem, too, is in a state of limbo that can be resolved only by choosing the axiomatic universe.

The result also helps to give us a broader understanding of learnability, Yehudayoff says. "This connection between compression and generalization is really fundamental if you want to understand learning."

Researchers have discovered a number of similarly 'undecidable' problems, says Peter O'Hearn, a computer scientist at University College London. In particular, following work by Gödel, Alan Turing — who co-founded the theory of algorithms — found a class of questions that no computer program can be guaranteed to answer in any finite number of steps.

But the undecidability in the latest results is "of a rare kind", and much more surprising, O'Hearn adds: it points to what Gödel found to be an intrinsic incompleteness in any mathematical language. The findings will probably be important for the theory of machine learning, he adds, although he is "not sure it will have much impact on the practice". ■