

SDS PODCAST

EPISODE 756

FIVE-MINUTE

FRIDAY:

ALPHAGEOMETRY:

AI IS SUDDENLY AS

CAPABLE AS THE

BRIGHTEST MATH

MINDS



(00:06):

This is Five-Minute Friday on AlphaGeometry.

(00:19):

Welcome back to The Super Data Science Podcast. I'm your host, Jon Krohn. As I've recently begun doing in some recent Five-Minute Friday episodes, let's kick things off with a couple of recent podcast reviews.

(00:30):

Eduardo Gonzalez, who works at the brewer Anheuser-Busch, he let us know that he's been promoted from Business Analyst to Senior Product Owner for Analytics. He says "This will be my first role stepping into a Data Science Team. [The] podcast was really helpful, helping me understand the vast problems DS is tackling." Nice, congrats on that promotion into your first data science role Eduardo.

(00:55):

And, another review, Kelly Adams, a data analyst at Golden Heart Games said: "Thanks for your show. It was one of the first data related podcasts I listened to and still one of my favorites." Great to hear that Kelly, thank you. Thanks for all the recent five-star ratings on Spotify, Apple Podcasts and all the other podcasting platforms out there. On Apple Podcasts, you can leave text feedback that I'd be delighted to read on air just like I've read the reviews today and that I assume helps more people discover the show. All right, let's jump right into the meat of the episode now.

(01:29):

Today, we're exploring a fascinating development from Google DeepMind, arguably the world's most prestigious AI lab. A couple of weeks ago, through a publication in the prestigious peer-reviewed journal Nature,



DeepMind introduced AlphaGeometry, a remarkable system with the ability to solve complex geometry problems, rivaling the capabilities of the world's brightest mathematical minds. So how complex are the problems that they were solving?

(01:54):

To make this comparison, DeepMind had their AlphaGeometry algorithm tackle geometry questions from the International Mathematical Olympiad. On a test of 30 of the latest Olympiad-level problems, AlphaGeometry solved 25 of them. How good is that? Well, on the same problems, the human Olympiad bronze medalists from the past two decades averaged 19.3 correct and silver medalists averaged 22.9. Only the Olympiad gold medalists narrowly edged out AlphaGeometry with an average of 25.9 problems solved compared to AlphaGeometry's 25. Not only does that mean that 25 is better than bronze medalists and silver medalists and comes very close to being as good as the gold medalists, so best geometric minds, pre-university geometric minds on the planet, that 25 problems absolutely crushes the previous machine state-of-the-art, which solved just 10 of the 30. So in terms of human intelligence on these complex geometry problems, AlphaGeometry is approaching the smartest geometric people on the planet even silver medalists in these competitions can compete, and they absolutely crushed the previous state-of-the-art, yeah, jumping from 10 in the previous state-of-the-art to 25 of the 30 questions solved successfully by AlphaGeometry.

(03:21):

So, now we know AlphaGeometry is super impressive, but how does it master deep mathematical reasoning like it does? Fundamentally, DeepMind chose to tackle challenging geometry problems because they require high levels of reasoning and logical deduction. Geometry, in particular, with its intricate diagrams and the necessity for creative problem-solving, presents a unique challenge to AI systems. Traditional



machine learning models falter when they are faced with the task of deciphering these geometric puzzles, primarily due to the complexities involved in translating human proofs into a format that machines can understand.

(03:57):

AlphaGeometry was designed specifically to answer the longstanding problem of teaching AI the art of geometric deduction. Unlike previous attempts, which depended on human proofs, human-generated proofs, the big leap here, maybe arguably, the biggest leap of all that DeepMind made here was that they used a novel approach of generating synthetic theorems and proofs. So this means that they weren't dependent on the human-generated proofs which obviously is going to be relative limited and very expensive data set to accumulate. So, instead of relying on human data, they simulated 100 million unique examples of geometric proofs, very effectively sidestepping the bottleneck of depending on human solutions to geometry problems.

(04:51):

Ok, so we've got abundant data, what about the model? AlphaGeometry is an innovative combination of a neural language model with a symbolic deduction engine. So let's break down those two components. The neural model, that means a neural network, an artificial neural network, like a deep learning system, it is akin to the Large-Language-Model architecture behind AI systems like ChatGPT, and this neural model serves as a system for providing intuitive guesses. So this is like the "fast thinking" that's described by the Nobel prize-winning economist Daniel Kahneman in his mega-bestselling book "Thinking, fast and slow". This fast-thinking LLM component, this neural model component of the AlphaGeometry system, predicts which geometric constructs, such as points, lines, or circles, might be useful in solving a particular problem. Then, after that, it's the second component in the system, the "symbolic deduction engine" that kicks in.



This “symbolic deduction engine” acts analogously to human slow, methodical thinking, it uses intuitive predictions from the fast, intuitive LLM component to logically deduce the steps required to arrive at the right solution.

(06:09):

This hybrid system, blending fast and slow thinking approaches together like in our own minds, has proven to be highly effective for geometry, evidently. Again, this allowed AlphaGeometry to solve 25 out of 30 recent Olympiad-level geometry problems, outperforming previous computational methods by a huge margin and showcasing a level of proficiency comparable to an International Mathematical Olympiad gold medalist. What’s even more impressive, that’s not all, is that the system solves these problems by generating human-readable proofs, and so this is a critical step towards making AI-generated solutions more accessible and understandable to humans.

(06:51):

The success of AlphaGeometry raises fascinating questions about the future of AI and its role in advancing human knowledge. The authors of the paper from DeepMind envision a future where AI can generalize across various mathematical fields, pushing the boundaries of what we know and how we solve problems. And their ambition doesn’t stop there: they envision their blend of fast and slow thinking to extend beyond mathematics, hinting at a future where AI could discover new knowledge across a spectrum of scientific disciplines.

(07:22):

The implications of AlphaGeometry are obviously then profound. By demonstrating that a hybrid of neural networks and symbolic reasoning can effectively tackle complex mathematical problems, DeepMind not only



advances the field of AI but also offers a glimpse into how future AI systems might learn, might reason, and might innovate. Very cool, particularly following so quickly on the heels of the publication of DeepMind's FunSearch algorithm in December, an algorithm for solving open mathematical problems using LLMs combined with an evolutionary learning approach that searches for functions, hence the name FunSearch, so this functions are written in computer code.

(08:02):

Unlike many other big LLM releases from the major AI research labs in recent years, you'll be delighted to know that DeepMind has fully open-sourced the code for both AlphaGeometry and FunSearch. This is presumably because there are fewer ethical concerns, at least at this time, of the misuse of math algorithms relative to natural language-generating models.

(08:22):

All right, that's it for today. I hope you found today's episode to be exhilarating, AI advances sure are moving mind-bogglingly quickly. Ok, until next time, keep on rockin' it out there and I'm looking forward to enjoying another round of the Super Data Science podcast with you very soon.