Stanford eCorner

The Future of Deep Learning

26-04-2023

URL: https://stvp.stanford.edu/clips/the-future-of-deep-learning

Ilya Sutskever, co-founder and chief scientist of OpenAI, says he expects deep learning to continue making progress, though the pace is different now. In the past, he says, progress could come from scaling, but now large training runs require more time and infrastructure. However, he believes improvement in many areas will still lead to robust progress.

Transcript

- I expect deep learning to continue to make progress.. 00:00:07,533 I expect that.. There was a period of time where a lot of progress came from scaling, and you saw that most in the most pronounced way in going from GPT-1 to GPT-3.. But things will change a little bit.. The reason that progress in scaling was so rapid is because people had all these data centers, which they weren't using for a single training run.. So by simply reallocating existing resources, you could make a lot of progress.. And it doesn't take that long necessarily to reallocate existing resources.. You just need to.... Someone just needs to decide to do so.. It is different now because the training runs are very big and the scaling is not going to be progressing as fast as it used to be because building data centers takes time..

But at the same time, I expect deep learning to continue to make progress in art from other places.. The deep learning stack is quite deep, and I expect that there will be improvements in many layers of the stack.. And together, they will still lead to progress being very robust.. And so if I had to guess, I'd imagine that there would be maybe, I'm certain we would discover new properties which are currently unknown in deep learning, and those properties will be utilized.. And I fully expect that the systems of five to 10 years from now will be much, much better than the ones they are we have right now.. But exactly how it's going to look like, I think it's a bit harder to answer.. It's a bit like.... It's because the improvements that there will be maybe small number of big improvements and also a large number of small improvements all integrated into a large complex engineering artifact.. - And can I ask? 00:02:07,410 Your co-founder, Sam Altman, has said that we've reached the limits of what we can achieve by scaling to larger language models.. Do you agree? And if so, then what is the next innovation frontier that you're focusing on, if that's the case? - Yeah, so..

00:02:26,707 I think maybe, I don't know exactly what he said, but maybe he meant something like that the age of easy scaling has ended or something like this.. Like, of course, of course the larger neural net will be better, but it will be a lot of effort and cost to do them.. But I think there will be lots of different frontiers.. And actually, to the question of how can one contribute in deep learning, identifying such a frontier, perhaps one that's been missed by others, is very fruitful.. - And can I go even just deeper on that? 00:03:00,090 Because I think there is this debate about vertical focus versus generalist training.. Do you think there's better performance that can be achieved in particular domains, such as law or medicine, by training with special datasets? Or is it likely that generalist training with all available data will be more beneficial? - So at some point, 00:03:23,790 we should absolutely expect specialist training to make a huge impact.. But the reason we do the generalist

training is just so that we can even reach the point where.... just so that we can reach the point where the neural network can even understand the questions that we are asking.. And only when it has a very robust understanding, only then we can go into specialist training and really benefit from it.. So yeah, I mean I think these are all fruitful directions...