Jay Yagnik

In his 15 years at Google, Jay Yagnik has led many foundational research efforts in machine learning and perception, computer vision, cyber security, quantum artificial intelligence (AI) and more. He has contributed significantly to the success of some of Google's most popular projects, like Google Photos, YouTube, Search and Maps.

An alumnus of the Indian Institute of Science (IISc), Bengaluru, Yagnik was back at his alma mater for a public talk on 4 December 2019. Following the talk, in a brief sit-down, he spoke about AI, his work at Google AI, the ethical and moral concerns of this technology, and the apocalyptic notions often associated with the predicted disruption.

Tell us a bit about yourself and your journey from being a student at IISc to Google?

It has been a good journey. Back when I was at IISc, I used to work in machine learning and computer vision. My thesis had an intersection of both, and I'd imagined working for some biometrics company, doing facial recognition and related things. Through a series of coincidences or [because of] people I met with an open mind, I ended up talking to Yahoo Labs at the time. I figured if I'm talking to one search company, why don't I talk to the other, and that's how I ended up talking to folks at Google. I walked out of those conversations with a very distinct impression: there was a lot of intellectual bandwidth. I was forced to think hard, even in real-time as the topics were evolving, and I took that as a positive sign for the work environment I would have if I were to join them. So I took a leap of faith and ended up making that move.

When I joined Google, the number of people who spoke [about] computer vision, machine learning and those domains, you could fit them all in one conference room, which we did back in the day. As we were starting on the early projects, I made a conscious choice to do projects that were different from the specific I had worked in. So some of the early projects I took on were in the video analysis domain, and those ended up panning out. As I gained more exposure to the Google environment, I saw clearly that we were sitting on some very unique assets. We were a search company, so we had all these web data, text data, we had Google Video and YouTube afterwards. I drafted a research agenda of building systems that look at images, read the surrounding text, look at videos and learn automatically, and that ended up being a good long-term trend. Over time, that led to some 18–20 projects that made many interesting things happen.

I then ventured outside [computer] vision into graph algorithms, into mainstream machine learning, and computer science theory. Every time I switched to a new domain, it gave me a chance to think afresh. Over time, [it] helped me to get a good top view of these technical fields and gave me this realization that we are solving roughly the same set of problems over and over in different disciplines with slightly different names and slightly different variants and environmental constraints.

Can you tell us what artificial intelligence is, perhaps by contrasting it with natural intelligence?

Actually, there is more than one definition of AI, but they all point at roughly the same direction. I guess in layman's terms, it is to build systems that exhibit behaviour that people would say is intelligent and is beyond a set of simple rules that are prescribed. If you go deeper than that, there are notions of what we call artificial general intelligence (AGI), which is to build systems that, when put in arbitrary environments, can exhibit intelligent behaviour given the constraints of the environment. They're not task-specific, but task-agnostic. We are still far from realizing those versions of AI. The things that are working well today are what we call task-specific AI, where you're building say, natural language understanding systems or machine perception systems or systems that are making recommendations or ranking or personalization of various forms. They end up being useful in practice, but are built for a [specific] task or a set of tasks.

So that's state-of-the-art today. When you talk about research, there's a lot of activity around AGI. There are teams that pursue the reinforcement learning view of the world, which is task-agnostic. You put the agent in an environment, and you give the agent reward signals, [and] it learns to gain more reward and [do] whatever it takes to get there. There are other approaches like extreme multitask learning – if you build one system that understands images and video, and also understands the text and is a dialogue system, eventually it is forced to learn intermediate representations that are shared across all of these, and that would lead to the notion of general intelligence over time. All of these are active research trends. But what you see in products and practical applications are more of the narrow AI because it is easier to integrate [it] into the context of a product which requires a specific question to be answered.

Can you give us a sense of some of the areas where AI is being used today and what sort of problems it is being used to address?

In any place where we are trying to associate some input with either a numerical value or a tag or a natural language outcome, [we] use machine learning and AI systems today. The examples I gave in the talk were in the context of AI for social good. For example, flood forecasting – predicting which specific terrain is likely to be flooded if a river overruns. Another would be in detecting diseases in plants, where you take a picture of the leaf, and categorize that picture into [either] a healthy leaf picture or one with a specific disease. If you look at Google Assistant, there you are providing a natural language interface. So your input is

raw audio samples which in the intermediate state get converted into text, and through a variety of other machine-learning transformations, this leads to a response in natural language, which is hopefully the answer to the intent you were expressing.

Let's say you have a problem, say detecting the likelihood of developing cancer based on MRI scans. When an AI system arrives at a solution, we don't have a sense of how it solves the problem – it is essentially a black box. How do we trust such a black box?

First of all, the AI system doesn't have to be a black box. That is often a stereotypical characterization, but AI systems can be made to work with people in the loop. We have many demonstrations of the same. If you look at our work in the health space, where we use medical records to predict the possibility of adverse outcomes in a hospital setting, the AI doesn't just predict the probability of an adverse outcome; it then goes back to the case history and picks out a small number of things so that the doctor can actually look at [and determine] which of them are statistically significant. So the way it manifests itself in practice is that the system says: 'Looks like there is a greater than 80% chance something might go wrong with this hospitalised patient, and from their entire lifelong case history, here are five things you should really look at.'

So, AI, in a way, is not replacing doctors but augmenting their ability to diagnose

Yes, it is both augmenting the person, and also scaling and leveraging their abilities. If you take our work on diabetic retinopathy, we find this to be the case pretty much. Expert ophthalmologists are rare [in many places]. So over two to three years, we managed to get that to work to a level where the accuracy of that system is on par with experts in the domain. When we look at countries like India or Indonesia, where we are actually using these systems, the ratio of people who need to be screened by ophthalmologists is skewed. So here's a way in which ophthalmologists can use their knowledge to look at cases which are borderline and require more nuanced expert opinion. And even when those cases come to the ophthalmologists, they don't

come as a black box; the system has already attended to parts of the image. It also explains the output that it is giving.

Alongside these medical use cases, you alluded to a few other applications where AI can have societal impacts. Can you tell us more about that?

Yes, if I stick to AI for social good, we've been doing some on-device machine learning. We open-sourced parts of it, which an NGO, Rainforest Connection, picked up. The NGO did this fascinating thing where it took old recycled Android phones, ran audio processing models on them, connected them to solar chargers and put them up on trees in the Amazon rainforest. These things look out for audio of bulldozers and chainsaws. And in real-time, they alert the authorities on illegal deforestation activities. These kinds of interventions are only possible with the technical capabilities that AI is uniquely unlocking.

In the Indian context, we've recently funded the Wadhwani AI Institute. I mentioned [in my talk] the project around the cassava plant imaging; they are doing something similar for agriculture in India. We've funded that through our Google.org or AI impact challenge grant. We're also doing work on the analysis of satellite imagery, which over time can lead to a better understanding of agriculture patterns, resource allocation and so on.

Another compelling use in the Indian context is this product called Google Lens, where you can point your camera at what you're looking at and it will bring up an experience that is relevant to the picture. People who cannot read can use Google lens and point at a piece of text. It will recognize the text, OCR it, and if necessary, translate it into a language that they understand and read it to them. This is now also allowing people who are illiterate to actually transact in the real world, because they are able to read and understand what's happening through this augmentation.

These are all the use cases where AI can be used for social good. But many commentators have also discussed the potential threats of AI – that we could lose control over the very things we have built. What are your thoughts about such concerns?

A lot of those narratives are coming from science fiction. What we find is that people tend to be bimodal when they make these assessments. For example, several people assume that if you achieve AGI, then it will have some version of a consciousness of its own. But if you look at the research, the more formal notions of AGI do not require these things to be present. It is unclear if they are necessary for intelligent behaviour. Really well-meaning people, because they are not practitioners in the field, can take two or three of these questionable assumptions and then jump to these conclusions. What we see in practice is that these things, when done with the right product mindset, can unlock some tremendously useful capabilities.

Barring apocalyptic notions of AI, there's also a moral concern. For instance, if a self-driving car gets into an accident with another car, one model of morality could make the car save as many lives as possible. But if you were driving your own car, your first instinct would be to save your life first and not necessarily worry about other lives.

If I were to be precise, I wouldn't use the term 'morality'. I would say ethics because morality is an overloaded word and depending on perspective and population group, it has different meanings. The question of ethics in AI is an important one, and at Google, we have written our AI principles. It is a collection of seven principles that practitioners need to be cognisant of in all of the macro decisions that they are making in building a system. There are things like fairness. [For instance, we could ask], 'Is the machine learning system that you're building producing equitable outcomes for different population groups?' And if you are careful algorithmically in building these systems, they can actually lead to more equitable outcomes.

So it is not just things that could potentially go wrong. There is a tremendous amount of opportunity to reduce societal bias and bring equitability in societal processes through this. Ethics as a field has had more than two centuries to develop. It has a lot of models to offer us in this context. We've also drawn red lines on some of the applications we will not pursue because they are not in line with these principles. For example, we have said that we will not pursue AI

applications in autonomous weaponry or things that are directly intended to cause harm in one form or another.

How far are we from achieving artificial general intelligence – the sort of intelligence we associate humans with?

It is hard to estimate a timeline for AGI because there is no one precise definition of it that everyone is striving towards. If you zoom out, there are, largely speaking, two types of AGI definitions or two camps within the AI community. There is the camp that looks at intelligent behaviour as an intelligent response to the environment. For example, the work from our DeepMind team, AlphaGo and all the gaming-related successes we've had, they all come from this worldview, that all you need is an agent that has exposure to the environment, and over time it learns to exhibit intelligent behaviour as measured by that environment. It doesn't say it needs to understand the text, or it needs to understand images or anything of that sort.

Then there's another camp which believes that building some underlying version of world knowledge that is grounded in some form is a prerequisite to achieving general intelligence. All the examples I gave on massively multitask learning or building systems that can read text, read images and understand audio at the same time, they're all explorations coming from this knowledge camp of AI. It remains to be seen over the next decade what ends up making more progress. AGI, at this point, is an interesting direction in which we're making progress. It is not a singular endpoint that everybody's aiming for.

Where do you think AI stands currently and where it will be in the next 20–25 years?

Take the evolution of photography as a technology. There have been several inflexion points, where photography merged with other disciplines. A lot of those inflexion points have to happen with AI, like AI and healthcare, AI and social sciences, AI and programming. I think those have to fully play out over the course of more than a decade. AI itself, in a technological sense, has to develop much more. Because if you look at the parts of AI that are really working well, it is the subset that we call supervised machine learning: you collect data, you then apply a class label to these data, and eventually train a machine learning system to map from input to the class label. That explains a vast number of use cases that are successful today. If we had this talk 10 or maybe 15 years from now, we would be at a point where AI and machine learning are more like a turnkey system that is not too data-heavy, and you can talk to it for a little bit, and it starts exhibiting intelligent behaviour. That would be a pipe dream today, but hopefully a reality a decade from now.

Google AI has recently started a lab in Bengaluru. Can you tell me more about the lab and what it intends to do going forward?

Yes, we recently started an AI lab in Bengaluru. Google has a presence in India and in Bengaluru, where lots of product development happens, but the emphasis of the AI lab is research. Our goal is to make this a world-class research lab, and it will be a peer to all of our other research outlets worldwide. Over time, it will also engage in some Google product development but driven by research activities.

(A version of this interview will also appear in *Connect*, the magazine of IISc.)

Pratik Pawar (S. Ramaseshan Fellow)
Current Science Association, Bengaluru 560 080, India.
e-mail: pratik@ias.ac.in