mint

From faulty cricket decisions to self-driving car deaths: The AI paradox

Devina Mehra | 4 December 2024

We discuss wrong cricketing decisions made by the system far longer than we do the more numerous mistakes made by human umpires.

SUMMARY

Human biases cause frequent failures but we expect technology to be entirely error-free. We should use AI to improve upon outcomes from human judgement, not reject it for imperfection.

A recent Primer on artificial intelligence (AI) in this newspaper mentioned all the ways AI could go wrong. Some of the issues related to AI taking over, like robots coming up with a language of their own. But there were also examples like self-driving cars causing accidents or AI giving wrong medical diagnoses.

Even in cricket, while rule-enforcing systems have been in place for a while, there are still occasional murmurs from players, commentators and spectators about the number of errors these make. People still talk about the wrong dismissal in 2014 of Pakistani opener Shan Masood in a Test match against New Zealand on account of a review call taken by the Decision Review System.

It is rather obvious that human beings will be reluctant to outsource decision-making to systems like this which make errors. This appears to be the right way to do things. But is it?

Let us do a thought experiment.

India has more than 150,000 people dying each year in road accidents—among the highest in the world even on a percentage-of-population basis.

Now suppose all vehicles are changed to self-driving ones and the death toll drops to 50,000.

What will newspaper headlines be?

Think carefully before you read any further. Will the headlines be, “Self driving vehicles reduce death toll by two thirds"? Or will they be, “Self driving vehicles kill 50,000 Indians every year." My guess is, it will be the latter.

This is the paradox!

My favourite academic Daniel Kahneman talks about this phenomenon in his book *Noise*, albeit not in the context of AI or self-driving cars.

One key takeaway: A well designed algorithm/rule-based system will almost always beat a so-called experienced expert in any area of human enterprise that requires judgement.

The reasons are simple. One, human beings are prone to biases, which are more systematic. For instance, biases for or against a race, caste, gender, etc. as well as others like loss aversion bias or hindsight bias. Two, having human decision-makers adds noise, which is random variability, to the mix.

Equally experienced experts in areas like judicial sentencing, insurance or investing will differ dramatically in making a judgement on the same issue and with exactly the same facts or information. The same person may come to divergent conclusions at different points in time. Tip: judges are more generous in granting bail or other reliefs after, rather than before, lunch.

This variability is noise, which can be reduced by having a proper system.

You can easily visualize this: Give the same company information to an array of stock market experts, analysts or fund managers, and each will have their own take on it.

We have all seen that skilled well-known doctors will often differ in their opinion both on the diagnosis of an ailment and the recommended treatment.

That, in short, is noise.

The book illustrates that in almost all human endeavours, right from reading mammograms to whether an accused should be granted bail, well-constructed algorithms consistently outperform human beings—that too humans who have a lot of experience and expertise—say, compared with judges with decades of experience.

So what’s the problem? It lies in how we judge the competing systems.

We intuitively know that human beings will make errors. For instance, doctors routinely misdiagnose ailments and human cricket umpires make mistakes. But we consciously or not expect a machine-led system, say an AI system, to be error free.

We are willing to ditch it at the first mistake: the first wrong diagnosis based on a mammogram or the first accused out on bail who commits a crime. Never mind that doctors and judges are also error-prone. In short, instead of evaluating whether the machine works better than human beings, we expect the machine system to be 'perfect'.

This is irrational. What we should be testing is whether the machine system improves on the alternative, rather than whether it is completely error-free on a standalone basis.

This usual pattern of thinking leads to wrong choices, as we may abandon a system or machine even if it is better than what was being done earlier.

That is why in the case of self-driving cars, we are not willing to live with a single fatality, even though human car drivers make many more errors and cause more deaths. It is also the reason why we discuss wrong cricketing decisions made by the system far longer than we do the more numerous mistakes made by human umpires.

The moral of the story is this: When it comes to evaluating alternative processes or systems, always pause and think, especially whether you are using the same yardstick to evaluate all the systems or have unrealistic expectations of one.

As an aside, this book is an endorsement of the path that we chose at First Global some years ago to put all our decades of research expertise into a system, which could then be applied on a bias- and noise-free basis across the whole universe of stocks. Of course, this system does not outperform in every single month or quarter—and that is not a bug!

Devina Mehra is chairperson, managing director and founder of First Global, an Indian and global asset management company, and author of the forthcoming 'Myths and Mantras: The Ultimate Investment Guide'. Her X handle is @devinamehra