

HO CHIN KUAN

educate
@thestar.com.my



Live&Learn

WELCOME to the future, where artificial intelligence (AI) systems augment, automate or replace human decision-making.

Imagine applying for a bank loan through an online system; you key in all pertinent information and almost instantaneously, the system informs you that you do not qualify for a loan.

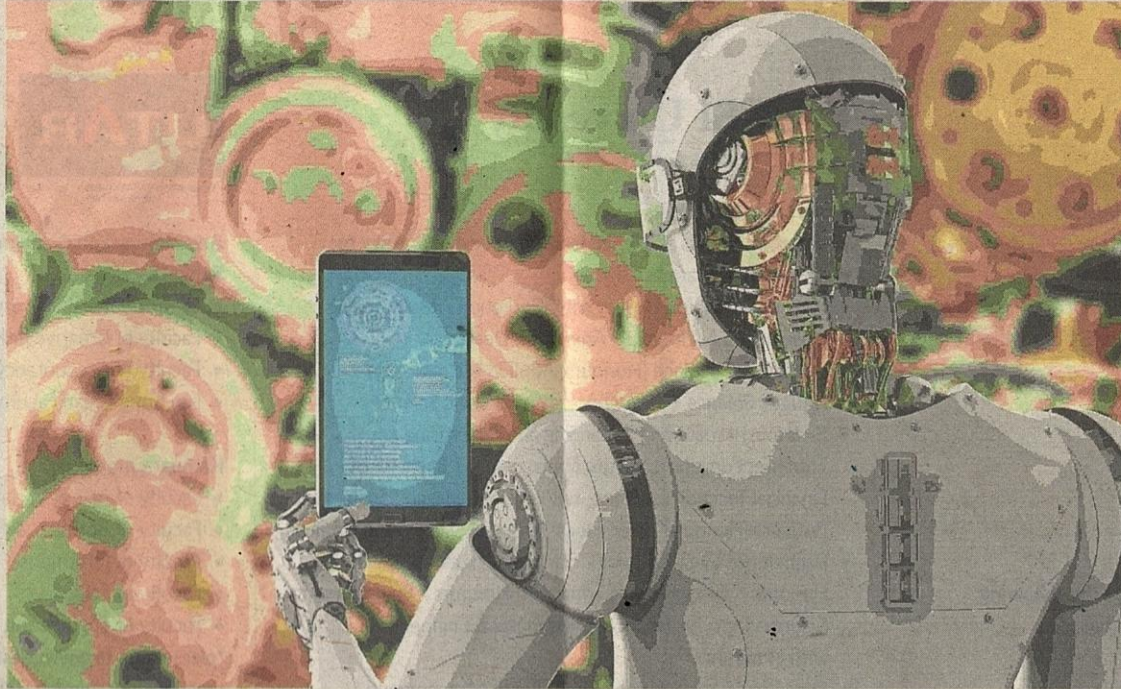
Coincidentally, you know that a friend with a profile very similar to yours got his loan approved by the same system.

Let's look at a second scenario: you decide to look for a new job – you send your resume to an online hiring system that immediately tells you you're not a right fit. A peer of yours who you think does not qualify has better luck.

The questions looming in your mind are – have I been fairly treated? How can I be sure that the AI system did not discriminate against me?

Can AI be fair?

Teach the leaders of tomorrow to create harm-free technology that will work for humanity



Humanising AI: Technology must embody the values we hold dear. – 123rf.com

DON'T MISS IT

TARCIan Run

Registration deadline:
Aug 5

> The much-anticipated TARCian Run to be held on Sept 16 is back for the sixth year running.

> The run will take place simultaneously across all Tunku Abdul Rahman University College (TAR UC) campuses in Kuala Lumpur, Penang, Johor, Perak, Pahang and Sabah.

> The event is jointly organised by TARCian Alumni Association (TAA) and TAR UC.

> The run aims to attract 10,000 runners in three categories: the 12km TARCian (Men), 12km TARCian (Women), and 7km "Fun Run".

> The event is open to all TARCians and their family members.

> All runners will get a free T-shirt. Those who complete the run within a stipulated time will also receive a medal.

> More than RM20,000

Human decision-making may sometimes be perceived as unfair, but shouldn't a computer with no human intervention produce fair decisions?

AI concerns building machines, hardware or software capable of performing tasks commonly requiring human intelligence. AI has experienced rapid growth in recent years due to the proliferation of computing power.

Today, AI is deployed for mission-critical applications, from hiring to disease diagnosis. Almost all AI applications we see now are a form of narrow AI, systems built for a specific task.

For example, a system built to navigate an autonomous vehicle will not be able to play a game of chess. The holy grail of AI research is to create artificial general intelligence. Such systems are self-aware, sapient and sentient. Think of the character Data in the *Star Trek: The Next Generation* series.

While waiting for artificial general intelligence to land, significant strides have been made in an area of AI called machine learning (ML). Voice-controlled personal assistants available in today's computing devices use some form of ML technology – take Siri, Cortana and Google

Assistant, for example.

While we may have read about the tremendous achievements of AI, such as the AlphaGo software from DeepMind beating the best human player at the game of Go, we have also seen reports raising grave concerns over social fairness and equity. Here are examples to illustrate these concerns:

- In *Machine Bias*, an online article dated May 23, 2016, ProPublica reported that software used across the United States to predict future criminals is biased against a particular segment of society.

- *The New York Times* published the article *Facial Recognition Is Accurate, if You're a White Guy* on Feb 9, 2018. It cited an M.I.T. Media Lab study which found that specific facial recognition software gives a 99% accuracy when an input photo is a white man. In comparison, up to 35% error was observed for darker-skinned women.

- More recently, a paper published in the research journal *Nature Machine Intelligence* dated June 17, 2021, reported that texts generated by the GPT-3 language model tend to associate people of a particular

faith with violence.

So apparently, ML suffers from biased decisions. We would have thought machines with no emotional or cultural attachments would do a better job. How did this happen? An ML system must learn from data to perform its task.

Data is generated from human activities and workflows. Our biases and cultural nuances are inherent in data. When we train AI systems to perform functions as we humans do, biases encoded in the data will influence the built models.

When applied, decisions by the models will likely reflect the bias in the data used for training. We enter a self-reinforcement loop where biases are amplified in new data.

Making ML systems work with the concept of fairness is challenging for several reasons.

One would need to represent the concept of fairness in a mathematical form that computers can understand.

This task is complex because the notion of fairness depends very much on our human value systems. Even if we can translate fairness requirements into mathematical terms, research has shown that some

fairness requirements are contradictory.

Engineers use a common technique to make ML systems “fairer”: de-bias data. A relatively straightforward way to do this is to omit sensitive attributes like race and gender.

Another method is to alter the outcomes of the ML system to align with the fairness objective. Some will see these actions as a form of affirmative action, drawing debate.

So, can machines be fair? In the larger sense, we ought to humanise AI technology. An excellent place to start is in higher education. We can find no shortage of technology-oriented programmes in our universities today.

We must teach our students, the leaders and shapers of tomorrow, to create technology that will work for humanity without harm or sacrificing the values that we hold dear.

To achieve this, there needs to be greater collaboration between the technical discipline and the social sciences. We need to break disciplinary silos.

For example, AI-related courses should be taught alongside social justice concepts; call this trustworthy AI. Let's look to a

worth of prizes are up for grabs.

> Besides engaging students in unity and solidarity, the other objective of the run is to raise funds for TAR UC. Since 2015, TAA has raised a total of RM200,000 for the TARC Education Fund.

Details:
www.tarc alumni.my

future where AI will seamlessly and safely augment our human ability.

Prof Dr Ho Chin Kuan is the vice chancellor at Asia Pacific University of Technology & Innovation (APU). He is also a fellow at the Overseas Chinese Development Research Center of the Yangtze Delta Region Institute of Tsinghua University, China. As an avid educator and researcher, his interests include data science, artificial intelligence, machine learning and complex systems. Prof Ho works with leading educators to co-build the future of EdTech. The views expressed here are the writer's own.