BIZCOMMUNITY

FaceApp and the dark side of AI

By Lee Naik

14 Aug 2019

Terminator and *The Matrix* got it wrong. We all thought the villains of the AI era would be killer robots or hostile programmes. Instead, they're face-ageing apps and YouTube creators with too much time on their hands.

Turns out the biggest problem with machine learning isn't the machines at all - it's the people behind them.

FaceApp may look like just another fun viral craze, but look closely at those terms and conditions and you'll realise you've signed away your image to a little-known Russian business forever.

The fears around FaceApp are just another reminder that technologies like AI aren't blank slates free from the intentions of their creators. It's no secret that, for all their potential to do good, algorithms are as prone to bias as the people who create them. Cognitive technologies are designed to imitate the learning process of the human mind - it's only natural that it should also emulate its flaws.

There's also another more sinister problem - what happens when you have bad actors exploiting those weaknesses or purposefully creating something intended to do harm?

Recently, YouTube has come under fire for pushing users towards harmful and extremist viewpoints over time. Then there's Twitter's much-publicised struggle with white supremacist accounts. We've gone from Tay, the racist chatbot, to psychopathic AIs and Russian hackers.

Manipulated algorithms

The problem is extremists know exactly how to manipulate search algorithms to expose their world views to more people. While legitimate organisations are getting to grips with 4IR, criminals have been comfortably using automation to commit fraud for years. And put the wrong people in charge of facial recognition technology and suddenly you've got a real threat to personal privacy.

In the wrong hands, machine learning can become a force for harm. It's up to the rest of us to make sure it lives up to its incredible potential to do good and solve problems.

Actions matter

We all remember Google's "don't be evil" motto, which it dropped in 2018 to great scepticism. What most people don't remember is the replacement line - "do the right thing".

It's a fitting change in the age of fake news and picture-perfect video forgery. Setting out to do no harm is all well and good, but it's not enough anymore. Organisations are waking up to the fact that the only way to prevent their platforms from being used to do harm is to proactively do good.

For some, this means fighting fire with fire. The UK government has a machine learning tool designed to automatically detect ISIS propaganda. Twitter and Google themselves are employing their AI tools to comb through and scrub their own data.


Lee Naik is the CEO of TransUnion Africa.

Many are also trying to define a universal set of ethics for machine learning and AI - an updated version of Asimov's Laws of Robotics you might remember from the movie *I, Robot*. The EU and the OECD are among those working towards a set of AI principles, and most major tech players have their own ethics boards.

But we shouldn't rely on the Googles, governments and Will Smiths of the world to do the heavy lifting. Every organisation has a role to play in shaping AI as a force for good. From the smallest chatbot to Watson itself, we should be working to create algorithms that offer transparency and accountability.

How to build an accountable AI

By their very nature, artificial intelligence technologies are somewhat unpredictable, but there are ways to make sure you're always on the right track. There are many strategies, guidelines and even Asimov's laws of robotics to lean on but for every organisation, I believe there are three key principles to throw into the mix:

- **The data is the key** - Any digital technology is only as good as its data, and this is even truer for artificial intelligences that teach themselves over time. Throw in enough flawed data early on and suddenly you've got a vicious cycle of bad assumptions multiplying other bad assumptions. Any AI is going to need a wide variety of reliable data streams to learn effectively.

- **Use purpose as your guiding light** - Ultimately, any algorithm needs a reason to learn in the first place, a clear guiding purpose. Without a strong purpose, it's easy for an ML application to go astray, and just as easy for the people behind the application to miss that it's happening.

- **Learn from your machines** - Finally, it's important not to get too focused on the 'machine' part that you forget about the 'learning'. The power of cognitive technologies isn't in how smart they are but in how much they enable us to embrace learning and continuous improvement as a way of life. Make sure your data and technology teams embody this, so that they are learning from the deviations and the unexpected detours of your AI applications.

Ultimately, the best way to extract real value and outcomes from technology is to make sure there are good people with good intent and accountability behind them. Because if you're not willing to take ownership, someone else will find a way.

ABOUT LEE NAIK

Lee Naik is the CEO of TransUnion Africa.

- NFTs and Web 3.0 to shape Africa's future in 2022 - 21 Jan 2022
- One app to rule them all: Rise of the super app - 8 Oct 2019
- FaceApp and the dark side of AI - 14 Aug 2019
- The digital identity question - 6 Mar 2019
- Lessons from Rio 2016 - technology versus human spirit - 25 Aug 2016

View my profile and articles...