Covid-19 & machine learning

With current clinical and computational ability, we should do better than during earlier pandemics

S ANANTHANARAYANAN

How the world is dealing with the pandemic has many dimensions. Apart from advanced understanding of the interior of cells and the mechanism of viral action, we now have sophisticated microscopic imaging and computing ability, and instant communication, for widespread cooperation.

The other modern tool that we have is machine learning – or the use of computers to analyse data, where the computers teach themselves to improve the accuracy of patterns in the data they discern. Machine learning uses the ability of computers to carry out massive computations to imitate the way neuron circuits in animal brains adapt and train themselves to sensitive pattern recognition.

A paper in the journal, *Nature Machine Intelligence*, by Li Yan, Hai-Tao Zhang, Jorge Goncalves, Yang Xiao, Maolin Wang, Yuqi Guo, Chuan Sun, Xiuchuan Tang, Liang Jing, Mingyang Zhang, Xiang Huang, Ying Xiao, Haosen Cao, Yanyan Chen, Tongxin Ren, Fang Wang, Yaru Xiao, Sufang Huang, Xi Tan, Niannian Huang, Bo Jiao, Cheng, Yong Zhang, Ailin Luo, Laurent Mombaerts, Junyang Jin, Zhiguo Cao, Shusheng Li, Hui Xu and Ye Yuan, from different departments in the Tongji Medical College, Schools of AI, Engineering and Information Science of the Huazhong and the Wuhan Universities of Science and Technology, Wuhan, China, Centre for System Biomedicine, Luxembourg and the University of Cambridge, describes a method of early and accurate assessment of the course a case of Covid-19 would take. Such an assessment helps optimise the use of available facilities by speedy segregation of persons testing positive, into groups that need different levels of care.

A simple application of machine learning is regression, or predictions based on past trends. An example, from an online presentation of machine learning by the Stanford University, is of estimating the price of a house from data of the floor area, number of rooms, bathroom, et al. The "learning data" is real information collected by a survey of houses that have been bought or sold.

Considering only the covered area, the data in respect of a sample of six houses could be like in Table 1. The same data is shown in the diagram, where the prices are plotted against the areas. The line that passes through the trend shown by the points could then be used to indicate nearly the correct price of a seventh house, either for the seller to decide what price to ask or for the buyer to decide what price is reasonable.

Machine learning uses a formal method to work things out from data like this. The price of a house is taken to be some base price plus an amount that depends on the area, like this: Price = base + rate x area.

Now, the idea is to discover the values of "base" and "rate" that best fit the data that we have. This is done by first working out a "cost function". The cost function is the difference of the price of each house, as shown by some assumed values of "base" and "rate", and the actual price. The total of this cost, for all the houses, gives us the cost function. This is when the computer gets active. It rapidly works out the different values of the cost function when the values of "base" and "rate" are varied. Different values are tried out, till we arrive at the lowest value for the cost function. These are then the values of "base" and "rate" that most closely match the available data.

The formula can then predict what the price may be of the seventh, eighth houses, and so on. When these houses are actually sold, and the real values are known, they can be added to the table and the estimation improved, till the formula becomes stable and reliable. is

This is when prices depend only on covered area. There would be other factors, of course, like the number of rooms, bathrooms, kind of material used, whether there is a parking, or a garden, and so on. With these factors added, the formula takes the form of,

Price = base + rate x area + factor.1 x feature.1 + factor.2 x feature.2 +...and so on.

The price can again be worked out by assuming values for the factors, and the best fit with actual data discovered by calculating the "cost function" and finding the optimum "factors".

Another kind of problem that machine learning deals with is of "classification". The example used in the Stanford study is of classifying tumours as malignant or benign. The example uses size of the tumour as the relevant feature. The data then consists of tumours of different sizes, and whether they turned out to be malignant or benign. Here, the answer that we are seeking is not one that can take all values, like the price, but just a Yes or a No – or a classification of tumours based on size. The data and graph would be like in Table 2.

In practice, of course, whether a tumour is malignant depends on many factors, like the features that decide the price of a house. The "value" derived from the complex data is expressed in a way that it is not a continuous number, but either "1" or "2", for "true" or "false" and a "cost function", of how far predictions are from facts, is similarly worked out, to arrive at a "decision boundary", to alert the physician if there are further tests she needs to do.

The animal brain also does complex classification, but the method is not to work out a "cost function", it is strengthen or eliminate responses to stimuli, depending on experience of outcome. For instance, if a bird finds that light colour and a grainy feel has been an edible tidbit on many occasions, the response of pecking is strengthened. But if the same colour with a shiny feel was a pebble, the bird learns not to peck.

The process is simulated in the computer, with the probability of a classification increased when combinations of features led to correct results, and the method is able to quickly become very expert – examples are of computers driving cars in traffic or of beating grandmasters at chess.

The group of scientists working at Wuhan have used methods like these to study a sample of 375 patients who tested positive for Covid-19. The initial symptom were fever, cough, fatigue and breathing problems. The 75 features considered included basic information, symptoms, blood samples and the results of laboratory tests, including liver function, kidney function, coagulation function, electrolytes and inflammatory factors. Against these features, recorded early in the infection, was the final result – recovery or mortality. And machine learning used the data to develop a scheme that identified the crucial biomarkers that were associated with the most serious prognosis. The scheme was then applied to 110 fresh instances of patients and the results were found to be more than 90 per cent accurate.

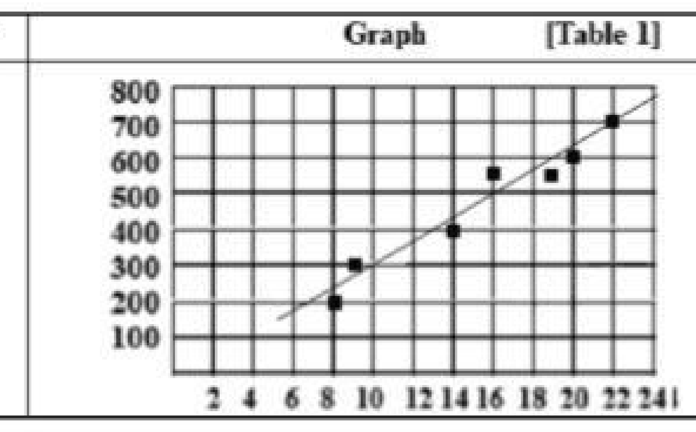
The results, in short, are that there are three vital factors – levels of lactic dehydrogenase (LDH), which reflects tissue damage; lymphocytes, or white blood cells, and high-sensitivity C-reactive protein, or hs-CRP, which reflects the state of inflammation. Identifying these factors gives the medical team advance pointers of what procedures to adopt, leading to both selection of the patients as well as giving selected patients the treatment that is most likely to help.

The significance of the work is twofold, the paper says. Apart from pointing out the high-risk factors. "It provides a simple and intuitive clinical test to precisely and quickly quantify the risk that the patient faces."

The writer can be contacted at response@simplescience.in

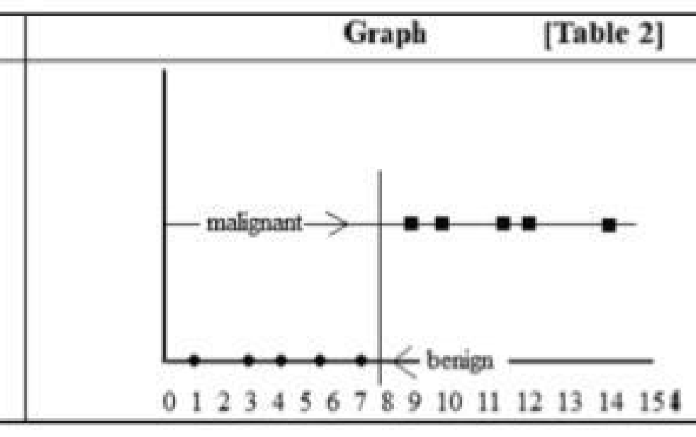
Graph [Table 1]

Area sq ft	Price Rs lakh
550	16.0
600	20.0
400	14.0.
550	20.0
450	14.0
700	22.0
200	8.0
300	9.0



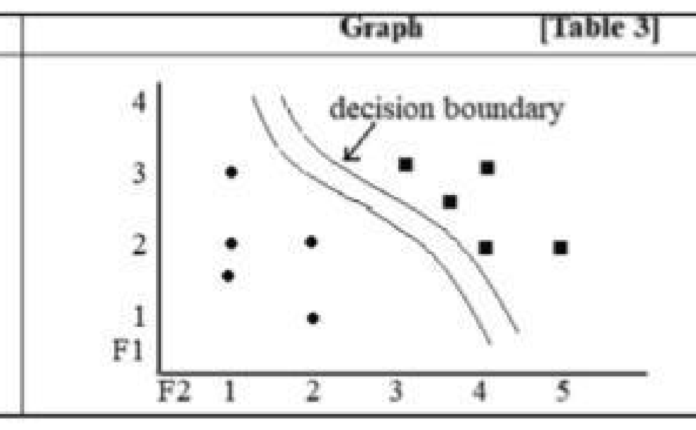
Graph [Table 2]

Size in mm	Whether malignant
3.0	N
7.0	N
11.5	Y
1.0	N
10.0	Y
12.0	Y
1.0	N
14.0	Y
4.0	N
9.0	Y



Graph [Table 3]

Features	Whether malignant
2.0-1.0	N
1.5-1.0	N
3.0-3.0	Y
2.5-3.5	Y
2.0-4.0	Y
1.0-2.0	N
2.0-5.0	Y
3.0-1.0	N
3.0-4.0	Y
2.0-2.0	N



Worthy of attention

Three videos of Unidentified Flying Objects were recently released by the Pentagon. But people don't seem as bothered or interested as they should be

ADAM DODD

On 27 April, the US Department of Defense issued a public statement authorising the release of three "UFO" videos taken by US Navy pilots.

The footage appears to depict airborne, heat-emitting objects with no visible wings, fuselage or exhaust, performing aerodynamically in ways that no known aircraft can achieve. The DoD doesn't use the terms "unidentified flying object" or "UFO" but does clearly state "the aerial phenomena observed in the videos remain characterized as 'unidentified'."

Thoughts about what UFOs are vary widely – from illusions to alien spacecraft. However, a workable, conservative definition is, "intelligently-controlled airborne objects not apparently made by humans".

Only a small fraction of UFO reports collected globally over the last seven decades seem to describe such objects, but the Navy footage appears to fit the bill. Whether such objects are vehicles of alien invasion or not, their mere presence would seem to indicate a national security threat, which is partly what makes the Pentagon's recent announcement so puzzling.

This is the first time the Pentagon has publicly confirmed the authenticity of UFO footage. It should have been a momentous announcement, but it seems to have barely moved the needle on the UFO controversy. Why?

The announcement is new, but the videos are not

The three grainy, monochrome infrared videos – one taken in November 2004, the other two in January 2015 – had already been leaked online, in 2007 and 2017, respectively. They also gained international attention after the *New York Times* published them as part of a December 2017 exposé on the Pentagon's secret UFO research programme, the so-called "Advanced Aerospace Threat Identification Program".

That programme was allegedly headed by Luis Elizondo, who claims to have been instrumental in the 2017 leaks, although his background has been credibly called into question. After resigning from the DoD, Elizondo immediately joined To the Stars Academy of Arts and Science, a UFO research collective founded by former Blink 182 frontman Tom DeLonge.

In September 2019, Joseph Gradisher, claiming the title of "spokesman for the deputy chief of naval operations for information warfare," confirmed the authenticity of all three videos in an email to a well-known UFO blog called *The Black Vault*. This development was quickly reported by the *Washington Post*.

The UFO footage in question, then, has appeared less like a shot out of the blue, and more like an echo in the night. Its gradual, staggered confirmation by the DoD mirrors the entrance of the footage itself into the public consciousness.

Whether this happened by accident or design, we may never know. As the technoculture critic Richard Thieme has astutely observed, "The UFO world is a hall of mirrors. The UFO world on the Internet is a simulation of a hall of mirrors."

Not ordinary, but not entirely invented

Despite the maddening refractions of the UFO rabbit hole, we can be certain of one thing. The modern figure of the UFO maintains an uneasy residence on "the margins of the real".

UFOs are clearly not ordinary objects, like rocks, chairs or smartphones. But neither are they utterly immaterial products of the cultural imagination, like werewolves, vampires or fairies. If, as historian of science M Norton Wise has argued, "To make something visible is to make it real, or to try to", then the question of whether UFOs exist or not largely hinges on debates about representation and authenticity. When it comes to phenomena that may not fit into our framework of what is real – phenomena like UFOs – what kind of representations of them will we regard as authentic?

More specifically, what would an authentic representation of a UFO look like? Who would have the authority to afford it that authenticity? And how would that authentication proceed?

What would 'legitimate' UFO footage look like?

In her widely influential 1977 polemic, *On Photography*, Susan Sontag observed "the images that have virtually unlimited authority in a modern society are mainly photographic images; and the scope of that authority stems from the properties peculiar to images taken by cameras".

Within this paradigm, even the poorest photograph is always more "legitimate" than the most refined and accurate painting. The Navy UFO footage is presented as something more than a photograph, however. It is offered as professional data, collected by highly skilled practitioners.

Even if we fail to fully understand everything on the plane's Advanced Targeting Forward-Looking Infrared display, or even how the video was made, it seems data-driven and authentic – an impression reiterated by the grainy, monochrome quality of the image itself. As observers, we are led to believe that, despite the somewhat visually disappointing resolution, we are watching authentic footage. In a way, the visual disappointment helps to qualify the videos as candidates for legitimacy.

Even though few of us know what such a video "should" look like, we assume that, since UFO encounters are spontaneous and surprising, footage is likely to be somewhat less than satisfactory.These expectations present a dilemma. If an image of a UFO is too clear it is likely to be read as obviously fake, but if it's too blurry it could be anything.

A superficial reading of the Navy UFO footage would likely lead to the latter evaluation. But given the nature of the footage (it is infrared, not technically photographic, so establishes the heat signature of the objects depicted), and the institutional context (the Pentagon is not known for producing and distributing fake UFO videos), it's hard to avoid concluding the footage shows genuine physical anomalies. If that's the case, it would be worthy of serious scientific and military attention, both of which currently seem absent.

'A hell of a video'

UFOs can be difficult and uncomfortable to think about. As I have argued elsewhere, one symptom of that difficulty is that individuals and institutions maintain their own ignorance of the situation. A persistent trope in Western UFO mythology is that every American President is briefed on the reality of the situation on taking office. The current President and commander-in-chief of the US Armed Forces, Donald Trump, commented on the recently released footage, "I just wonder if it's real. That's a hell of a video."

It was a rare unifying statement from a notoriously divisive and antagonistic President, perhaps encapsulating the most likely public reaction to this latest installment in the UFO mystery – just wonder.

The writer is a tutor at University of Queensland, Australia. This article first appeared on www.theconversation.com

PLUS POINTS

Live fast, die young

Size, safety and parenting all have an impact on how quickly a species of bird matures, according to new research from the UK that could help scientists understand and predict how animals will respond to climate breakdown and the destruction of habitats.

The team of scientists has studied thousands of species of birds to understand why there is so much diversity in the length of time they take to grow from a fertilised egg to an independent adult. The research, published in *Nature Communications*, is the first study to consider the importance of lifestyle and environmental factors alongside evolutionary history and body size to explain the variation.

All organisms face a trade-off between reproducing and surviving and they solve this problem in different ways. The team found that bird species with a "live fast die young" strategy develop quicker, allowing them to maximise the number of offspring they can produce in the short time they have available. Findings showed that birds that breed and live in safer environments with fewer predators typically took longer to develop, possibly because they can afford to spend longer in a vulnerable state. They also found that migratory birds develop much quicker, which may ensure they are ready to return to their winter habitats at the end of the summer. As expected, the research showed that bigger birds took longer to develop -- but even among birds of a similar size there was variation in development times.

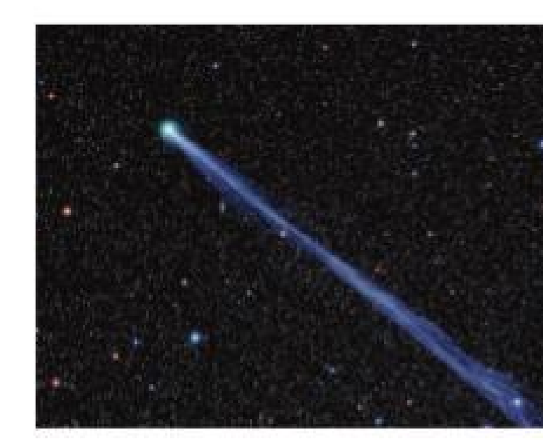
Chris Cooney, from the University of Sheffield's department of animal and plant sciences and lead author of the research, said, "The amount of time it takes for a fertilised egg to develop into a fully grown adult varies hugely across the animal kingdom. For instance, it takes an elephant almost 10 years to reach independence, whereas a fruit fly is fully grown after only a matter of days.

"This extraordinary diversity is also encapsulated within birds, where albatrosses can take almost a year to develop from an embryo to an independent adult, but a typical garden songbird takes little more than a month. We found that certain aspects of a species' lifestyle and environment are important in explaining how long they take to develop."

Alison Wright, co-author of the research from the University, said, "Our study on birds gives us some clues about the type of factors that may be important in other species. However, it may be that different factors are important for determining development length in other animal groups.

"The next step is therefore to address these questions using data that covers the breadth of the animal kingdom – from fish to mammals to insects – to gain an even broader insight into the factors shaping these fundamental differences across species."

Swan in the sky



A comet could be visible with the naked eye as it flies past Earth soon. The object, known as Comet Swan, was discovered by an instrument floating in space – Nasa and the European Space Agency's Solar and Heliospheric Observatory, or Soho, satellite.

As it gets closer, the comet should be visible in the Southern Hemisphere just before sunrise, without any equipment. The show could be seen in the coming weeks, becoming most clearly visible at the end of May and beginning of June.

Comet Swan came closest to Earth on 13 May, when it swung by around 53 million miles away. It will carry on to get closer to the Sun on 27 May, before sailing back off through the Solar System.

As comets get nearer to the Sun, and the temperature gets hotter, they tend to heat up and start shedding material in a dust trail that can be visible in images. The ice, dust and rock that makes up a comet can then break up or become more visible, and it is difficult to know how any given object will behave in the circumstances.

–The independent