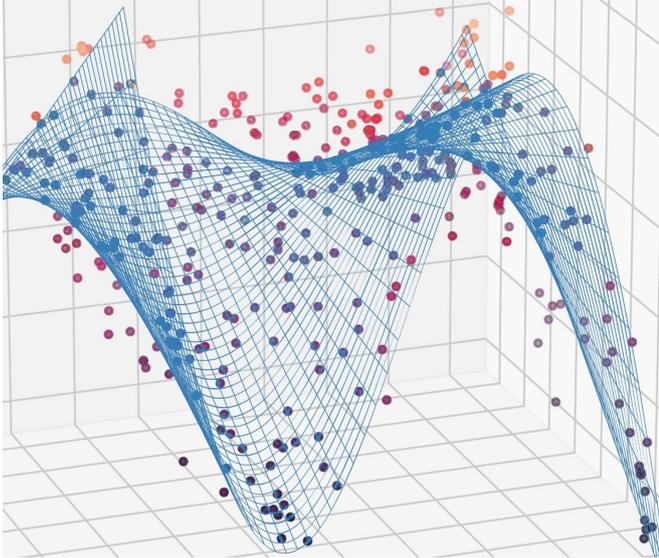# The most human algorithm

31 January 2020

From data (points) given to the scientific robot, it is able to find the law that best describes them (blue surface) and their exact mathematical expression. Credit: URV

It is now possible to predict who the best candidate for receiving an organ transplant is, know whether clients of a bank will return the loans they request, choose the films that best coincide with the interests of consumers, or even select someone's ideal partner. Mathematical algorithms constantly analyse millions of items of data, identify patterns and make predictions about all areas of life. But in most cases, the results give little more than a closed prediction that cannot be interpreted and which is often affected by biases in the original data.

Now, a team from the research group SEES:lab of the Department of Chemical Engineering of the Universitat Rovira I Virgili and ICREA has made a breakthrough with the development of a new algorithm that makes more accurate predictions and generates mathematical models that also make it possible to understand these predictions. The results of this research have just been published in the journal *Science Advances*.

"The aim of our study was to create what is known as a scientific robot, an algorithm that can apply the knowledge and expertise that a researcher has to interpret data," explains Marta Sales-Pardo, one of the authors of the paper. The results provided by the algorithm are characterised by the fact that they are interpretable. "It is as if someone had drawn up a law or a theory on the system that is being studied. The algorithm gives you the mathematical relations between the variables it has analysed and it does so completely independently," adds Roger Guimerà, an ICREA researcher from the same group.

When a company has an enormous amount of data that it wishes to exploit, it can do so by employing someone to try various models, propose formulas and find which one works best by carrying out experiments to validate them. This will lead to a mathematical formula that makes it possible to model the system but it involves a considerable investment in time and money.

Another possibility is to find a specialist in machine learning, a scientific discipline in the field of artificial intelligence that creates systems that identify complex patterns in enormous data sets, learn automatically and produce a "black-box" model that can make predictions. However, these systems provide no other information and if the prediction fails it is impossible know where the error lies and what needs to be done to prevent it.

The algorithm developed at the URV takes the best of the two cases: it processes the data automatically, quickly and reliably, as the machine learning system does, and it also produces a result that is an interpretable model.

The algorithm can be used to analyse and interpret data from any discipline in a process that is much more agile and efficient than those in existence to date. But the real added value is the information that the system provides. "In medicine, for example, if you have to take a decision based on data it is very important to understand why each decision

has been taken and the risk of making a mistake," explains Guimerà. "Although the algorithm has also shown that it is highly accurate, the most important thing is that you can understand the results because you have built a machine scientist that, with no previous knowledge, can take a set of data and develop a theory that solves the problem posed," adds Ignasi Reichardt, another researcher on the team.

In this study, the [algorithm](#) has been applied to a fundamental problem of fluid physics with the collaboration of the research group Experimentation, Computation and Modelling in Fluid Mechanics and Turbulence of the URV's Department of Mechanical Engineering.

More information: "A Bayesian machine scientist to aid in the solution of challenging scientific problems" *Science Advances* (2020). [DOI: 10.1126/sciadv.aav6971](#) , https://advances.science mag.org/content/6/5/eaav6971

Provided by Universitat Rovira i Virgili