Study shows how machine learning could predict rare disastrous events, like earthquakes or pandemics

December 19 2022

Credit: Pixabay/CC0 Public Domain

When it comes to predicting disasters brought on by extreme events

(think earthquakes, pandemics or "rogue waves" that could destroy coastal structures), computational modeling faces an almost insurmountable challenge: Statistically speaking, these events are so rare that there's just not enough data on them to use predictive models to accurately forecast when they'll happen next.

But a team of researchers from Brown University and Massachusetts Institute of Technology say it doesn't have to be that way.

In a [new study](#) in *Nature Computational Science*, the scientists describe how they combined statistical algorithms—which need less data to make accurate, efficient predictions—with a powerful machine learning technique developed at Brown and trained it to predict scenarios, probabilities and sometimes even the timeline of rare events despite the lack of historical record on them.

Doing so, the research team found that this new framework can provide a way to circumvent the need for massive amounts of data that are traditionally needed for these kinds of computations, instead essentially boiling down the grand challenge of predicting rare events to a matter of quality over quantity.

"You have to realize that these are stochastic events," said George Karniadakis, a professor of applied mathematics and engineering at Brown and a study author. "An outburst of pandemic like COVID-19, environmental disaster in the Gulf of Mexico, an earthquake, huge wildfires in California, a 30-meter wave that capsizes a ship—these are rare events and because they are rare, we don't have a lot of historical data. We don't have enough samples from the past to predict them further into the future. The question that we tackle in the paper is: What is the best possible data that we can use to minimize the number of data points we need?"

The researchers found the answer in a sequential sampling technique called active learning. These types of statistical algorithms are not only able to analyze data input into them, but more importantly, they can learn from the information to label new relevant data points that are equally or even more important to the outcome that's being calculated. At the most basic level, they allow more to be done with less.

That's critical to the [machine learning](link) model the researchers used in the study. Called DeepOnet, the model is a type of artificial neural network, which uses interconnected nodes in successive layers that roughly mimic the connections made by neurons in the human brain. DeepOnet is known as a deep neural operator. It's more advanced and powerful than typical artificial neural networks because it's actually two neural networks in one, processing data in two parallel networks. This allows it to analyze giant sets of data and scenarios at breakneck speed to spit out equally massive sets of probabilities once it learns what it's looking for.

The bottleneck with this powerful tool, especially as it relates to rare events, is that deep neural operators need tons of data to be trained to make calculations that are effective and accurate.

In the paper, the research team shows that combined with active learning techniques, the DeepOnet model can get trained on what parameters or precursors to look for that lead up to the disastrous event someone is analyzing, even when there are not many data points.

"The thrust is not to take every possible data and put it into the system, but to proactively look for events that will signify the rare events," Karniadakis said. "We may not have many examples of the real event, but we may have those precursors. Through mathematics, we identify them, which together with real events will help us to train this data-hungry operator."

In the paper, the researchers apply the approach to pinpointing parameters and different ranges of probabilities for dangerous spikes during a pandemic, finding and predicting rogue waves, and estimating when a ship will crack in half due to stress. For example, with rogue waves—ones that are greater than twice the size of surrounding waves—the researchers found they could discover and quantify when rogue waves will form by looking at probable wave conditions that nonlinearly interact over time, leading to waves sometimes three times their original size.

The researchers found their new method outperformed more traditional modeling efforts, and they believe it presents a framework that can efficiently discover and predict all kinds of rare events.

In the paper, the research team outlines how scientists should design future experiments so that they can minimize costs and increase the forecasting accuracy. Karniadakis, for example, is already working with environmental scientists to use the novel method to forecast climate events, such as hurricanes.

The study was led by Ethan Pickering and Themistoklis Sapsis from MIT. DeepOnet was introduced in 2019 by Karniadakis and other Brown researchers. They are currently seeking a patent for the technology.

Provided by Brown University