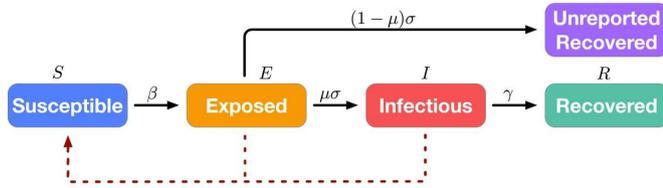


Machine-learning model is helping CDC predict spread of COVID-19

19 May 2020, by Christine Wei-Li Lee



A diagram of the machine-learning model. One feature of SuEIR is that people in both “exposed” and “infectious” can infect people in “susceptible,” as indicated by the dashed arrows. Credit: Samueli School of Engineering

A machine-learning model developed at the UCLA Samueli School of Engineering is helping the Centers for Disease Control and Prevention predict the spread of COVID-19.

The model was created by a team led by Quanquan Gu, a UCLA assistant professor of computer science, and it is now one of 13 models that feed into a [COVID-19 Forecast Hub](#) at the University of Massachusetts Amherst. Data from that hub, in turn, feeds into the CDC's [online forecasts](#) for how the disease might continue to spread.

Gu said his model is more accurate than most others because it does not rely only on confirmed COVID-19 cases and fatalities. It is epidemiology-driven and is one of only two models in the hub that use [machine learning](#).

The model's name, [UCLA-SuEIR](#), is derived from the five types of observed and inferred COVID-19 data that factor into its projections—the number of cases categorized as susceptible, unreported, exposed, infectious and recovered.

The UCLA model is unique because it doesn't simply fit the current curve, which is based only on

reported cases. Rather, it infers the number of untested and unreported cases from the model's data analysis and uses those inferences to predict how quickly the disease will spread. This is called an "epidemic model" because it takes into account the various factors that affect the rate of disease spread.

UCLA-SuEIR produces state- and county-level models based on the numbers of fatalities and confirmed [cases reported by the New York Times](#), and national models based on data reported by Johns Hopkins University.

The University of Massachusetts added UCLA's model to its hub on May 6 after Gu sent details about his work to UMass biostatistics professor Nicolas Reich, the hub's project lead. Gu's team had noticed that several models in the hub were producing varying predictions, mostly based on curve-fitting models.

"Without any epidemic modeling, the projection by the curve-fitting model is very misleading since it only depends on the observed data pattern but ignores the underlying epidemic dynamics that drive the data," Gu said.

The UCLA team checks its model's accuracy by making a prediction one week in advance of future confirmed cases, death and recovered cases, then verifying it against the actual reported data. The model's machine-learning algorithm enables Gu to train a new prototype in less than five seconds and allows the team to update its model on a daily basis, which is more efficient than other models. Gu's team has actually created a total of 232 sub-models in all—one for the U.S. overall, as well as one for each state and 181 for any counties with more than 1,000 confirmed cases.

Gu said the UCLA model has been consistently the most accurate in the Massachusetts hub in predicting death counts for the U.S. and most

states, and that it is among the top three models that best match their predictions with the actual number of reported deaths nationwide.

The work is much more than a mathematical exercise. "Our model can help measure the effectiveness of government policies, such as social distancing, stay-at-home orders, the use of face masks and covers, or self-quarantine, as well as predicting possible resurges in cases as states reopen," Gu said.

The model also could be used to evaluate whether a region is testing enough people, which can help officials understand whether more testing is needed.

According to the team's projections, the number of COVID-19 cases will peak June 1 for the U.S. overall, while California will reach its apex July 1, and Los Angeles cases will peak June 7.

Gu's team is working on refining the [model](#) to take into account county-level data from hospital and intensive care units, which could yield predictions that authorities could use to better allocate resources such as health care workers, personal protective equipment and ventilators.

Provided by University of California, Los Angeles

APA citation: Machine-learning model is helping CDC predict spread of COVID-19 (2020, May 19) retrieved 17 January 2021 from <https://medicalxpress.com/news/2020-05-machine-learning-cdc-covid-.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.