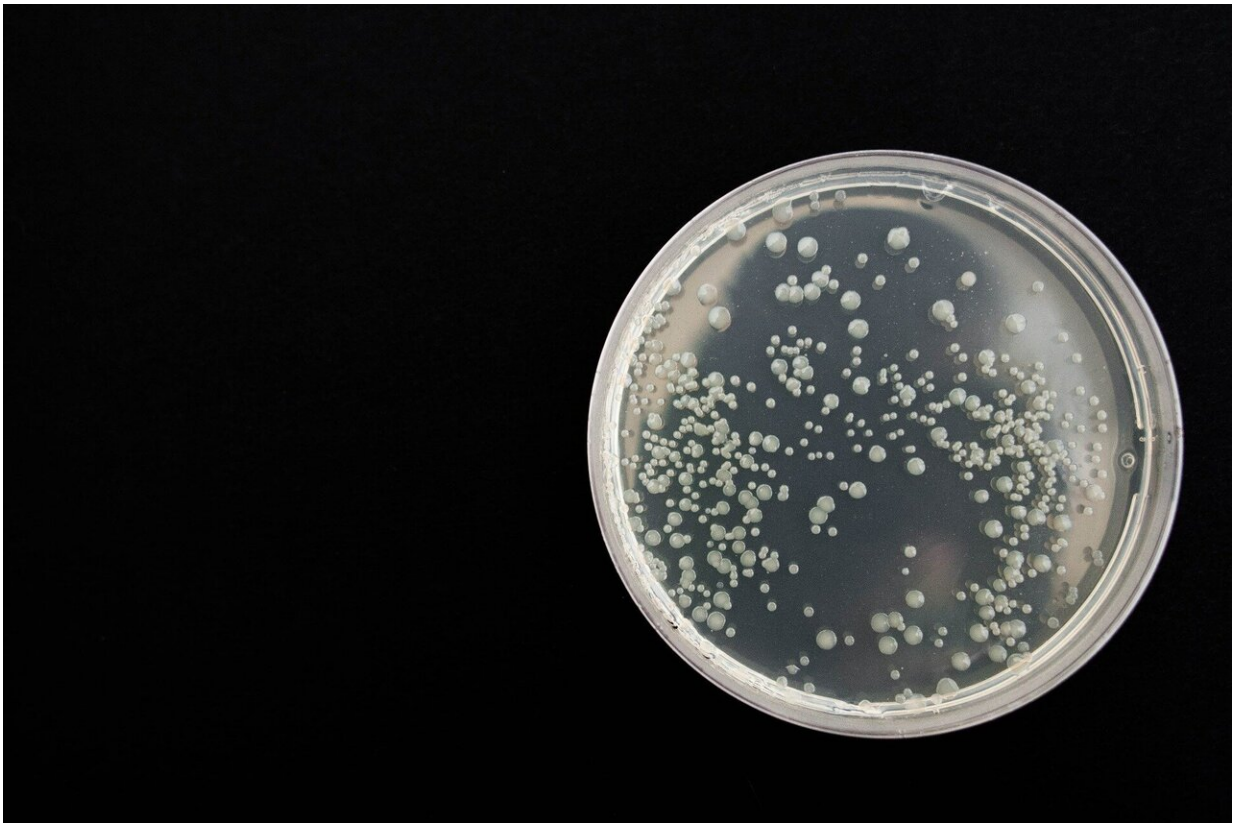


Machine learning models predict COVID-19 impact in smaller cities

April 17 2020, by Mary L. Martialay



Credit: CC0 Public Domain

According to a robust machine learning model that can predict pandemic impact even in smaller cities, with 75% of the population in the Capital Region in New York remaining at home, the COVID-19 pandemic will

peak locally in the second half of May. If the rate of people staying home drops to 50%, it will peak in early June. Rensselaer Polytechnic Institute researcher Malik Magdon-Ismael tailored the models he is developing to work with sparse data points, like those available during the early phase in a pandemic or in smaller cities, which ordinarily make trend-spotting difficult.

"There are no simple, robust, general tools that, for example, officials in Albany could use to make projections," said Magdon-Ismael, a professor of computer science, and expert in [machine learning](#), [data mining](#), and pattern recognition. "These models show that the projections vary enormously from one [city](#) to another. This knowledge could relieve some of the uncertainty that is around in developing policy."

Using county data available through the New York State Department of Health and Mental Hygiene, Magdon-Ismael has developed models that can predict local aspects of the [pandemic](#) such as the rate of infections over time, the infectious force of the pandemic, the rate at which mild infections become serious, and estimates for asymptomatic infections. The research model is ongoing work and, given the time-sensitive nature of the work, earlier versions have been released on the arXiv preprint server, which is moderated but not peer-reviewed.

His model for the Capital Region—which incorporates the data from Albany, Rensselaer, Saratoga, and Schenectady counties up to April 10—uses a total at-risk population of 855,000 to estimate that daily confirmed infections will peak at 1,490 on June 8 with 50% staying at home, or 750 on May 28 with 75% staying at home. The number of infections would total 58,000 or 29,000, respectively. Confirmed infections as of April 10 are approximately 1,000 and the model estimates 14,000 asymptomatic cases at that time.

Modeling smaller cities with machine learning is a challenge in that few

data points are available and updated less frequently than the picture of the nation as a whole or an epicenter like New York City. Generic machine learning operating on such data would likely produce inaccurate predictions. To compensate, Magdon-Ismail focuses on simple models and uses "robust" algorithms that incorporate solutions beyond that of the mathematical ideal.

"The machine gives you the [model](#) that best fits the data, but it turns out the best is usually a very fragile principle. There are lots of different models, lots of different explanations that are essentially as good," Magdon-Ismail said. "To make the output robust, we consider the collection of models that have near-optimal levels of consistency with the data. I find a variety of models that fit the data, and then I use all of those models together to predict."

Magdon-Ismail said producing similar models for other small cities in New York state would be as easy as "running the numbers."

In an earlier effort, also published online in arXiv, Magdon-Ismail tested his approach on data from the very beginning of the pandemic in the United States. With so few infections reported from January 20 to March 14, the early data was similarly as sparse as that available in small cities. Early data provided another insight, in that it offered a look at what the virus would do if unchecked.

"Early data is captured in the analogy: if you want to learn about a lion, you don't observe the lion in the zoo, you have to observe the lion on the savannah," Madgon-Ismail said. "And basically what that means is early dynamics of the pandemic. Nobody really knows what's going on, nobody really knows whether it's serious, so nobody's really done anything. And that's where you see how it will really behave."

More information: Machine Learning the Phenomenology of

COVID-19 From Early Infection Dynamics. arxiv.org/abs/2003.07602

Provided by Rensselaer Polytechnic Institute

Citation: Machine learning models predict COVID-19 impact in smaller cities (2020, April 17)
retrieved 3 May 2024 from

<https://medicalxpress.com/news/2020-04-machine-covid-impact-smaller-cities.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.