



Los Alamos AI model wins flu forecasting challenge

October 22, 2019

LOS ALAMOS, N.M., Oct. 22, 2019—A probabilistic artificial intelligence computer model developed at Los Alamos National Laboratory provided the most accurate state, national and regional forecasts of the flu in 2018, beating 23 other teams in the Centers for Disease Control and Prevention’s [FluSight Challenge](#). The CDC announced the results last week.

“Accurately forecasting diseases is similar to weather forecasting in that you need to feed computer models large amounts of data so they can ‘learn’ trends,” said Dave Osthus, a statistician at Los Alamos and developer of the computer model, Dante. “But it’s very different because disease spread depends on daily choices humans make in their behavior—such as travel, hand-washing, riding public transportation, interacting with the healthcare system, among other things. Those are very difficult to predict.”

The FluSight Challenge aims to improve accurate flu forecasting by challenging scientific institutions to develop predictive computer models. During the 2018-2019 flu season, 24 different teams participated in the flu forecasting initiative, each submitting 38 different weekly forecasts.

Dante proved more successful than the other models in predicting the timing, peak and short-term intensity of the unfolding flu season. Unlike other models, Dante is a multi-scale model, meaning it combines national, regional and state flu data. By averaging the trends across those different geographies, it uses information from individual states to improve other states’ forecasts.

Each week from mid-October to mid-May, Osthus submitted a file to the CDC that described Dante’s forecasts for the entire flu season. “Submitting each week of the season allows forecasters to update their forecasts in light of current data—similar to how, for instance, hurricane forecasts are updated as the hurricane is unfolding,” he said.

New data for the flu season are collected each week and integrated into the forecasting models. Dante proved particularly useful for forecasting at the local level, something that is, according to Osthus, “accompanied with significant data challenges.”

For this flu season, Osthus plans to submit Dante+, an updated version of Dante that will include internet-based “nowcasting,” which develops and uses a model that maps Google search traffic for flu-related terms onto official flu activity data.

As for what Osthus predicts for this year’s flu season, it’s hard to say. “Flu forecasts this early in the season are marked by significant uncertainty,” he said. “The flu season

doesn't usually start to reveal itself until after Thanksgiving. There is nothing, at this point, to suggest a highly unusual flu season, meaning it is likely to peak between mid-December and late March. As far as the intensity of the flu season, however, it's just too early to tell."

Kelly Moran (a Ph.D. student at Duke University and, at the time, a visiting guest student scientist at Los Alamos) contributed to the validation of Dante. The second-place model, DBM+, was also developed at Los Alamos with the help of Reid Priedhorsky, Ashlynn Daughton (a Ph.D. student at University of Colorado Boulder), Sara Del Valle and Jim Gattiker. The [Dante paper](#) can be viewed online.

Caption for image below: *Influenza-like illness (ILI) activity is highly spatially variable, with higher than typical levels of flu activity (pink) concentrated around the Gulf of Mexico, and typical (white) to below typical (green) ILI levels seen throughout the rest of the country. The spatial variability illustrates the challenge and importance of jointly modeling ILI for forecasting.*

Los Alamos National Laboratory

www.lanl.gov

(505) 667-7000

Los Alamos, NM

Managed by Triad National Security, LLC for the U.S Department of Energy's NNSA

