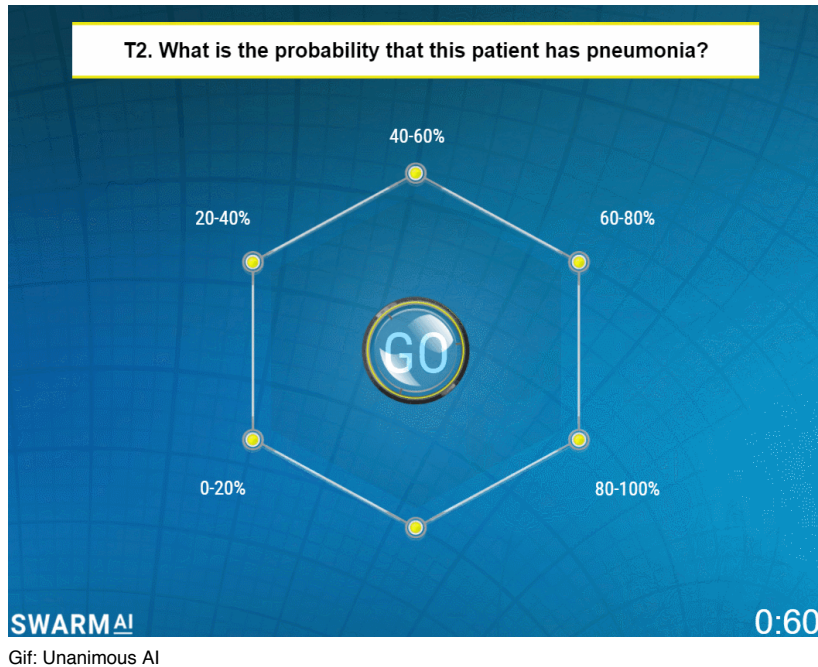


13 Sep 2018 | 17:54 GMT

AI-Human “Hive Mind” Diagnoses Pneumonia

A small group of doctors moderated by AI algorithms made a more accurate diagnosis than individual physicians or AI alone

By **Megan Scudellari** (/author/megan-scudellari)



First, it correctly predicted the top four finishers at the Kentucky Derby. Then, it was better at picking Academy Award winners than professional movie critics—three years in a row. The cherry on top was when it prophesied that the Chicago Cubs would end a 108-year dry spell by winning the 2016 World Series—four months before the Cubs were even in the playoffs. (They did.)

Now, this AI-powered predictive technology is turning its attention to an area where it could do some real good—diagnosing medical conditions.

In a [study \(https://cdn.ymaws.com/siim.org/resource/resmgr/mimi18/abstracts/18paper4-Halabi.pdf\)](https://cdn.ymaws.com/siim.org/resource/resmgr/mimi18/abstracts/18paper4-Halabi.pdf) presented on Monday at the SIIM Conference on Machine Intelligence in Medical Imaging in San Francisco, Stanford University doctors showed that eight radiologists interacting through [Unanimous AI \(https://unanimous.ai/\)](https://unanimous.ai/)'s “swarm intelligence” technology were better at diagnosing pneumonia from chest X-rays than individual doctors or a machine-learning program alone.

“It went really well,” says [Matthew Lungren \(https://profiles.stanford.edu/matthew-lungren\)](https://profiles.stanford.edu/matthew-lungren), a pediatric radiologist at Stanford University Medical School, coauthor on the paper and one of the eight participants. “Before, we had to show [an X-ray] to multiple people separately and then figure out statistical ways to bring their answers to one consensus. This is a much more efficient and, frankly, more evidence-based way to do that.”

It was a small study, but the findings suggest that instead of [replacing doctors \(/static/ai-vs-doctors\)](/static/ai-vs-doctors), AI algorithms might work best [alongside them \(/the-human-os/biomedical/diagnostics/ai-diagnostics-move-into-the-clinic\)](/the-human-os/biomedical/diagnostics/ai-diagnostics-move-into-the-clinic) in health care.

“We shouldn’t throw away human knowledge, wisdom, and experience,” says [Louis Rosenberg](https://unanimous.ai/staff-item/louis-rosenberg-phd/) (<https://unanimous.ai/staff-item/louis-rosenberg-phd/>), CEO and founder of Unanimous AI. “Instead, let’s look at how we can use AI to leverage those things.”

The company’s technology—a combination of AI algorithms and real-time human input—has also made [headlines](https://unanimous.ai/press-releases/) (<https://unanimous.ai/press-releases/>) by correctly predicting Trump’s approval ratings, TIME’s Person of the Year, and the exact final score of Super Bowl 51, among others.

The current study is the company’s first foray into medicine.

Pneumonia is a particularly tricky disease to diagnose on X-rays alone because it looks like a lot of other illnesses. In the current study, eight radiologists in different locations sat in front of their computers and analyzed 50 chest X-rays from an open source data set. Each doctor was asked to predict how likely it was that the patient had pneumonia, based on the X-ray.

But this was not crowdsourcing—each doctor did not simply respond with a “yes” or a “no.” Instead, using the [Swarm AI system](https://unanimous.ai/what-is-si/) (<https://unanimous.ai/what-is-si/>)—modeled on the collective decision-making process of honeybee swarms—each doctor controlled a small magnet icon that enabled them to push the group consensus toward their opinion. Every X-ray was examined in real-time with the other doctors simultaneously contributing opinions.

As the doctors weighed in, AI algorithms monitored the behavior of each participant, inferring how strongly each felt about their choice based on the relative motions of their icon over time. Someone who holds out longer on one choice, for example, may be expressing a stronger sentiment than someone who switches opinion quickly or several times.

“To really find the optimal solution, it’s not enough to just know what their opinions are, one really needs to know their varying levels of confidence,” says Rosenberg.

The algorithms then combined those preferences into a specific choice. Each deliberation took between 15 to 60 seconds, and the doctors diagnosed all 50 X-rays in about 90 minutes, says Rosenberg.

In the end, the Swarm AI system was 33 percent more accurate at correctly classifying patients than individual practitioners, and 22 percent more accurate than a Stanford machine-learning program called [CheXNet](https://stanfordmlgroup.github.io/projects/chexnet/) (<https://stanfordmlgroup.github.io/projects/chexnet/>). Last year, CheXNet [beat radiologists at diagnosing pneumonia](https://arxiv.org/abs/1711.05225) (<https://arxiv.org/abs/1711.05225>) from X-rays.

The Swarm AI technology is unlikely to be used by radiologists for the hundreds of chest X-rays that cross their desks daily, says Lungren, but it could be especially useful in two key situations. First, it would be “insanely invaluable” in situations where international experts are asked to weigh in on difficult cases, he says.

Second, the technology enables doctors to each have an equal chance to influence a diagnosis. When a group of doctors meets to discuss a difficult case, which is common in large hospitals, some of the smartest people in the room may be introverts and their voices might not be heard, says Lungren. Swarm AI takes politics and personalities out of the process.

“The best way to get multiple humans to agree on something, so far, for us, has been the swarm,” says Lungren.

The team now plans to conduct a larger study using actual patient cases at the Stanford University Medical Center.

The Human OS Newsletter

Biomedical engineering in a biweekly newsletter. Expert insights into wearable sensors, big data analytics, and implanted devices for personalized medicine.

About the Human OS blog

IEEE Spectrum's biomedical engineering blog, featuring the wearable sensors, big data analytics, and implanted devices that enable new ventures in personalized medicine.

Eliza Strickland (/author/strickland-eliza), Editor

Emily Waltz (/author/emily-waltz), Contributor

Megan Scudellari (/author/megan-scudellari), Contributor

Subscribe to RSS Feed (<https://spectrum.ieee.org/rss/blog/the-human-os/fulltext>)