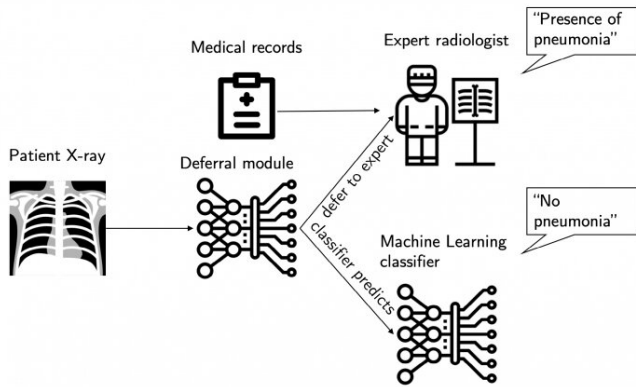


# An automated health care system that understands when to step in

3 August 2020, by Adam Conner-Simons



The system either queries the expert to diagnose the patient based on their X-ray and medical records, or looks at the X-ray to make the diagnosis itself. Credit: MIT CSAIL

In recent years, entire industries have popped up that rely on the delicate interplay between human workers and automated software. Companies like Facebook work to keep hateful and violent content off their platforms using a combination of automated filtering and human moderators. In the medical field, researchers at MIT and elsewhere have used machine learning to help radiologists better detect different forms of cancer.

What can be tricky about these hybrid approaches is understanding when to rely on the expertise of people versus programs. This isn't always merely a question of who does a task "better"; indeed, if a person has limited bandwidth, the system may have to be trained to minimize how often it asks for help.

To tackle this complex issue, researchers from MIT's Computer Science and Artificial Intelligence Lab (CSAIL) have developed a [machine learning system](#) that can either make a prediction about a task, or defer the decision to an [expert](#). Most

importantly, it can adapt when and how often it defers to its human collaborator, based on factors such as its teammate's availability and level of experience.

The team trained the system on multiple tasks, including looking at chest X-rays to diagnose specific conditions such as atelectasis (lung collapse) and cardiomegaly (an enlarged heart). In the case of cardiomegaly, they found that their human-AI hybrid model performed 8 percent better than either could on their own (based on AU-ROC scores).

"In medical environments where doctors don't have many extra cycles, it's not the best use of their time to have them look at every single data point from a given patient's file," says Ph.D. student Hussein Mozannar, lead author with David Sontag, the Von Helmholtz Associate Professor of Medical Engineering in the Department of Electrical Engineering and Computer Science, of a new paper about the system that was recently presented at the International Conference of Machine Learning. "In that sort of scenario, it's important for the system to be especially sensitive to their time and only ask for their help when absolutely necessary."

The system has two parts: a "classifier" that can predict a certain subset of tasks, and a "rejector" that decides whether a given task should be handled by either its own classifier or the human expert.

Through experiments on tasks in [medical diagnosis](#) and text/image classification, the team showed that their approach not only achieves better accuracy than baselines, but does so with a lower computational cost and with far fewer training data samples.

"Our algorithms allow you to optimize for whatever choice you want, whether that's the specific

prediction accuracy or the cost of the expert's time and effort," says Sontag, who is also a member of MIT's Institute for Medical Engineering and Science. "Moreover, by interpreting the learned rejector, the system provides insights into how experts make decisions, and in which settings AI may be more appropriate, or vice-versa."

The system's particular ability to help detect offensive text and images could also have interesting implications for content moderation. Mozanner suggests that it could be used at companies like Facebook in conjunction with a team of human moderators. (He is hopeful that such systems could minimize the amount of hateful or traumatic posts that human moderators have to review every day.)

Sontag clarified that the team has not yet tested the system with human experts, but instead developed a series of "synthetic experts" so that they could tweak parameters such as experience and availability. In order to work with a new expert it's never seen before, the system would need some minimal onboarding to get trained on the person's particular strengths and weaknesses.

In future work, the team plans to test their approach with real human experts, such as radiologists for X-ray diagnosis. They will also explore how to develop systems that can learn from biased expert data, as well as systems that can work with—and defer to—several experts at once. For example, Sontag imagines a hospital scenario where the system could collaborate with different radiologists who are more experienced with different patient populations.

"There are many obstacles that understandably prohibit full automation in clinical settings, including issues of trust and accountability," says Sontag. "We hope that our method will inspire machine learning practitioners to get more creative in integrating real-time human expertise into their algorithms."

**More information:** Mozannar et al., Consistent Estimators for Learning to Defer to an Expert. arXiv:2006.01862 [cs.LG]. [arxiv.org/abs/2006.01862](https://arxiv.org/abs/2006.01862)

*This story is republished courtesy of MIT News ([web.mit.edu/newsoffice/](http://web.mit.edu/newsoffice/)), a popular site that covers news about MIT research, innovation and teaching.*

Provided by Massachusetts Institute of Technology

APA citation: An automated health care system that understands when to step in (2020, August 3)  
retrieved 19 October 2021 from <https://techxplore.com/news/2020-08-automated-health.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.*