Machines Watch You Surf the Web
Machine learning systems control targeted ads and much more.

by
Austin Milt, Christine Dumoulin, Jonathan Reyles, and Stephanie Rickett
Doctoral Students
NSF IGERT SCALE-IT (Scalable Computing and Leading Edge Innovative Technologies)
University of Tennessee
Revised 11/12/10

There's something coming through the Wi-Fi that you didn't ask for. It's watching what you do. It's there when you open those "private" emails or post on a friend's wall. It talks to you, even if you don't realize it. And what's worse, you're talking back.

But no one is hiding behind your computer screen, secretly tracking your movements through the net. That task has been left to the computers and a technique called machine learning.

"[Machine learning] affects the kind of Internet advertising you see on Amazon.com, but most people don't realize that [it] is the underlying technology," says Dr. Cecily Heiner, who has worked on educational tools involving machine learning, at the University of Utah.

So if you've noticed that some ads seem to know what you've been up to, you aren't paranoid. You're onto something.

Many of the ads you see online are displayed by machine learning systems, commonly termed agents, whose purpose is to "learn" how people use the Internet. They compare your activity with everyone else's to determine whether you're more likely to respond to an ad for body wash or one for *sexy* body wash. If the agent gets it right and you click on an ad, then the ad publisher (think Google, Yahoo,or Facebook) charges a fee to the advertiser. Advertisers are willing to pay a lot to get their ads shown

(upwards of $70 per click for some Google AdSense Keywords!), giving publishers plenty of reason to invest in machine learning research.

Whether you find targeted ads enticing or annoying, machine learning isn't just for the Internet. "We are surrounded by computer-based systems that impact our everyday lives," says Dr. Itamar Arel, director of the Machine Intelligence Lab at the University of Tennessee.

Machine learning techniques are currently used in automobile navigation systems, noise-cancelling headsets, and red-eye reduction in cameras. Arel focuses on machine learning algorithms that aid doctors in interpreting medical images. He hopes to develop machines that learn just as well if not better than us humans. According to Arel, a good radiologist needs to look at about 1000 cases of benign and malignant growths before becoming proficient. "We're trying to mimic that same capability. [Given a large number of examples,] can we teach the system to learn what to look for by itself?" As computer power increases, these same machine learning agents may be able to diagnose cancers more quickly and with less training than even the best radiologists.

Machine learning systems don't come out of the box knowing about your health or your favorite book. Instead, they begin much like us: cute and stupid. Like school children learning arithmetic from a teacher, some agents can use simple examples and feedback to learn how to approach more complicated information. Others learn by experience, storing the consequences of their "decisions" and drawing upon them when faced with similar situations. Either way,
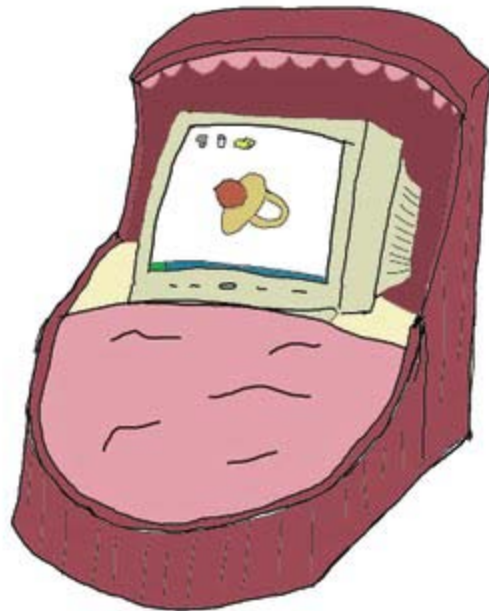
these machines can catch on to patterns in the data they're given. When it comes to placing advertisements, this allows them to quickly home in on your likes and dislikes so your next click is their doing.

Despite its ubiquity, machine learning has limitations. Because much of machine learning attempts to mimic human learning, our incomplete understanding of the human brain means that teaching a machine to "learn to learn" is easier said than done. Unlike us, machine learning programs can only deal with a limited range of information. The software that decides which advertisements you see cannot drive your noise-cancelling headphones or spot a problem on an X-ray of your head. And, if given too much information too early, a machine may get caught up in the details and not learn to catch patterns at all.

Image-interpreting machines like Arel's cannot process entire pictures at once, but have to wade through them pixel by pixel. Whereas you could take in the richness of a picture at a glance, computers prefer the thousand words. As scientists continue to learn about human behavior, researchers like Arel will continue to translate those results into machines that respond to the world more like we do.

Online advertising, medicine and other current uses are only the tip of the iceberg. "I suspect that in 10 years, there will be at least one [machine learning] component in almost any consumer electronics product," says Arel.

Heiner agrees: "There are some aspects of our life that we already seem perfectly content to turn over to machines … and I suspect that as we continue to improve technology, we will continue to allow it to control more aspects of our lives." Whether that means our cars will drive themselves or human doctors will become obsolete is ultimately up to advancements in hardware and us as users of machine learning technology.

Machine learning is here to stay, as a boost to our electronics or possibly something more.

"I don't think we will allow ourselves to fall into robotic servitude," says Heiner. But Arel isn't so sure: "[Past a certain point], it would be difficult to predict the future. I'd like to hope we'll manage to benefit from such a technology rather than have it lead to the destruction of the human race."