

Machine learning techniques generate clinical labels of medical scans

31 January 2018

Researchers used machine learning techniques, including natural language processing algorithms, to identify clinical concepts in radiologist reports for CT scans, according to a study conducted at the Icahn School of Medicine at Mount Sinai and published today in the journal *Radiology*. The technology is an important first step in the development of artificial intelligence that could interpret scans and diagnose conditions.

From an ATM reading handwriting on a check to Facebook suggesting a photo tag for a friend, computer vision powered by [artificial intelligence](#) is increasingly common in daily life. Artificial intelligence could one day help radiologists interpret X-rays, computed tomography (CT) scans, and [magnetic resonance imaging](#) (MRI) studies. But for the technology to be effective in the medical arena, computer software must be "taught" the difference between a normal study and abnormal findings.

This study aimed to train this technology how to understand text reports written by radiologists. Researchers created a series of algorithms to teach the computer clusters of phrases. Examples of terminology included words like phospholipid, heartburn, and colonoscopy.

Researchers trained the computer software using 96,303 radiologist reports associated with head CT scans performed at The Mount Sinai Hospital and Mount Sinai Queens between 2010 and 2016. To characterize the "lexical complexity" of radiologist reports, researchers calculated metrics that reflected the variety of language used in these reports and compared these to other large collections of text: thousands of books, Reuters news stories, inpatient physician notes, and Amazon product reviews.

"The language used in radiology has a natural structure, which makes it amenable to machine learning," says senior author Eric Oermann, MD,

Instructor in the Department of Neurosurgery at the Icahn School of Medicine at Mount Sinai. "Machine learning models built upon massive radiological text datasets can facilitate the training of future artificial intelligence-based systems for analyzing radiological images."

Deep learning describes a subcategory of machine learning that uses multiple layers of neural networks (computer systems that learn progressively) to perform inference, requiring large amounts of training data to achieve high accuracy. Techniques used in this study led to an accuracy of 91 percent, demonstrating that it is possible to automatically identify concepts in text from the complex domain of radiology.

"The ultimate goal is to create algorithms that help doctors accurately diagnose patients," says first author John Zech, a medical student at the Icahn School of Medicine at Mount Sinai. "Deep learning has many potential applications in radiology—triaging to identify studies that require immediate evaluation, flagging abnormal parts of cross-sectional imaging for further review, characterizing masses concerning for malignancy—and those applications will require many labeled training examples."

"Research like this turns big data into useful data and is the critical first step in harnessing the power of artificial intelligence to help patients," says study co-author Joshua Bederson, MD, Professor and System Chair for the Department of Neurosurgery at Mount Sinai Health System and Clinical Director of the Neurosurgery Simulation Core.

Provided by The Mount Sinai Hospital

APA citation: Machine learning techniques generate clinical labels of medical scans (2018, January 31) retrieved 18 September 2019 from <https://medicalxpress.com/news/2018-01-machine-techniques-clinical-medical-scans.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.