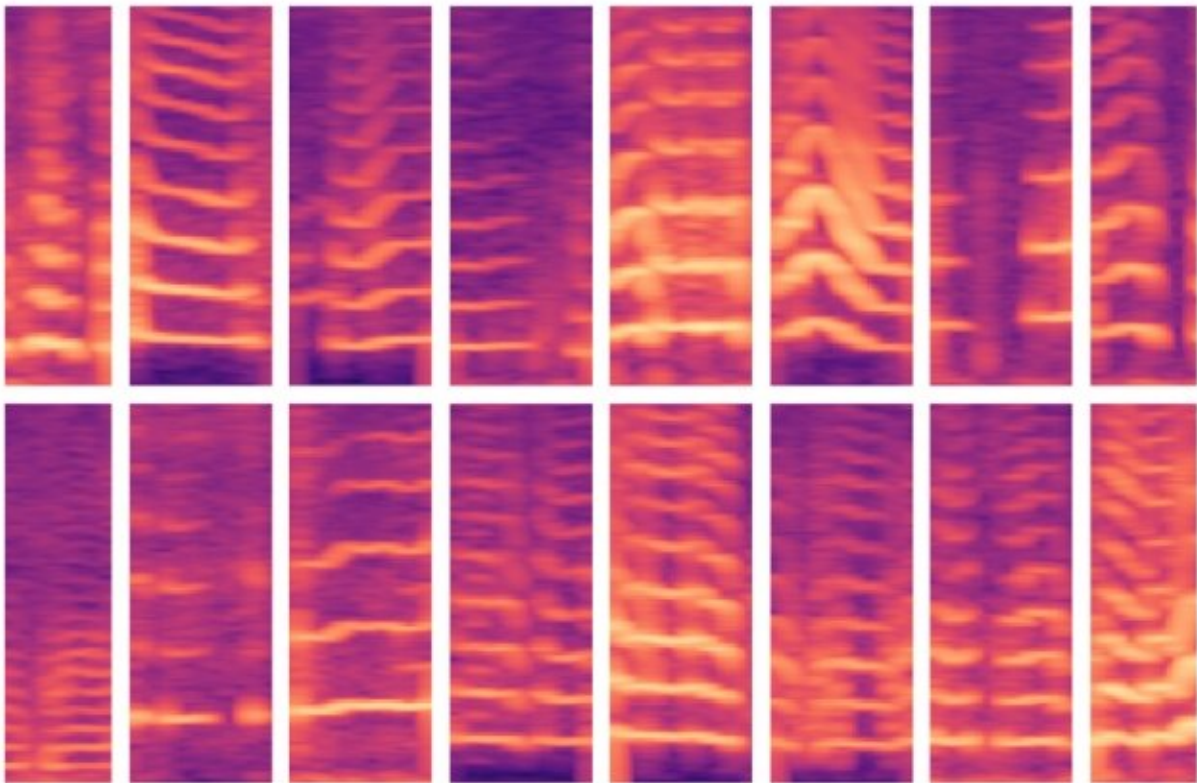


Automating artificial intelligence for medical decision-making

August 6 2019, by Rob Matheson



A new MIT-developed model automates a critical step in using AI for medical decision making, where experts usually identify important features in massive patient datasets by hand. The model was able to automatically identify voicing patterns of people with vocal cord nodules (shown here) and, in turn, use those features to predict which people do and don't have the disorder. Credit: Massachusetts Institute of Technology

MIT computer scientists are hoping to accelerate the use of artificial intelligence to improve medical decision-making, by automating a key step that's usually done by hand—and that's becoming more laborious as certain datasets grow ever-larger.

The field of predictive analytics holds increasing promise for helping clinicians diagnose and treat patients. Machine-learning models can be trained to find patterns in [patient data](#) to aid in sepsis care, design safer chemotherapy regimens, and predict a patient's risk of having breast cancer or dying in the ICU, to name just a few examples.

Typically, training datasets consist of many sick and healthy subjects, but with relatively little data for each subject. Experts must then find just those aspects—or "[features](#)"—in the datasets that will be important for making predictions.

This "feature engineering" can be a laborious and expensive process. But it's becoming even more challenging with the rise of wearable sensors, because researchers can more easily monitor patients' biometrics over long periods, tracking sleeping patterns, gait, and voice activity, for example. After only a week's worth of monitoring, experts could have several billion data samples for each subject.

In a paper being presented at the Machine Learning for Healthcare conference this week, MIT researchers demonstrate a [model](#) that automatically learns features predictive of vocal cord disorders. The features come from a dataset of about 100 subjects, each with about a week's worth of voice-monitoring data and several billion samples—in other words, a small number of subjects and a large amount of data per subject. The dataset contain signals captured from a little accelerometer sensor mounted on subjects' necks.

In experiments, the model used features automatically extracted from

these data to classify, with high accuracy, patients with and without vocal cord nodules. These are lesions that develop in the larynx, often because of patterns of voice misuse such as belting out songs or yelling. Importantly, the model accomplished this task without a large set of hand-labeled data.

"It's becoming increasingly easy to collect long time-series datasets. But you have physicians that need to apply their knowledge to labeling the dataset," says lead author Jose Javier Gonzalez Ortiz, a Ph.D. student in the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL). "We want to remove that manual part for the experts and offload all feature engineering to a [machine-learning model](#)."

The model can be adapted to learn patterns of any disease or condition. But the ability to detect the daily voice-usage patterns associated with vocal cord nodules is an important step in developing improved methods to prevent, diagnose, and treat the disorder, the researchers say. That could include designing new ways to identify and alert people to potentially damaging vocal behaviors.

Joining Gonzalez Ortiz on the paper is John Guttag, the Dugald C. Jackson Professor of Computer Science and Electrical Engineering and head of CSAIL's Data Driven Inference Group; Robert Hillman, Jarrad Van Stan, and Daryush Mehta, all of Massachusetts General Hospital's Center for Laryngeal Surgery and Voice Rehabilitation; and Marzyeh Ghassemi, an assistant professor of computer science and medicine at the University of Toronto.

Forced feature-learning

For years, the MIT researchers have worked with the Center for Laryngeal Surgery and Voice Rehabilitation to develop and analyze data from a sensor to track subject voice usage during all waking hours. The

sensor is an accelerometer with a node that sticks to the neck and is connected to a smartphone. As the person talks, the smartphone gathers data from the displacements in the accelerometer.

In their work, the researchers collected a week's worth of this data—called "time-series" data—from 104 subjects, half of whom were diagnosed with vocal cord nodules. For each patient, there was also a matching control, meaning a healthy subject of similar age, sex, occupation, and other factors.

Traditionally, experts would need to manually identify features that may be useful for a model to detect various diseases or conditions. That helps prevent a common machine-learning problem in health care: overfitting. That's when, in training, a model "memorizes" subject data instead of learning just the clinically relevant features. In testing, those models often fail to discern similar patterns in previously unseen subjects.

"Instead of learning features that are clinically significant, a model sees patterns and says, "This is Sarah, and I know Sarah is healthy, and this is Peter, who has a vocal cord nodule." So, it's just memorizing patterns of subjects. Then, when it sees data from Andrew, which has a new vocal usage pattern, it can't figure out if those patterns match a classification," Gonzalez Ortiz says.

The main challenge, then, was preventing overfitting while automating manual feature engineering. To that end, the researchers forced the model to learn features without subject information. For their task, that meant capturing all moments when subjects speak and the intensity of their voices.

As their model crawls through a subject's data, it's programmed to locate voicing segments, which comprise only roughly 10 percent of the data. For each of these voicing windows, the model computes a spectrogram,

a visual representation of the spectrum of frequencies varying over time, which is often used for speech processing tasks. The spectrograms are then stored as large matrices of thousands of values.

But those matrices are huge and difficult to process. So, an autoencoder—a neural network optimized to generate efficient data encodings from large amounts of data—first compresses the spectrogram into an encoding of 30 values. It then decompresses that encoding into a separate spectrogram.

Basically, the model must ensure that the decompressed spectrogram closely resembles the original spectrogram input. In doing so, it's forced to learn the compressed representation of every spectrogram segment input over each subject's entire time-series data. The compressed representations are the features that help train [machine-learning](#) models to make predictions.

Mapping normal and abnormal features

In training, the model learns to map those features to "patients" or "controls." Patients will have more voicing patterns than will controls. In testing on previously unseen subjects, the model similarly condenses all spectrogram segments into a reduced set of features. Then, it's majority rules: If the subject has mostly abnormal voicing segments, they're classified as patients; if they have mostly normal ones, they're classified as controls.

In experiments, the model performed as accurately as state-of-the-art models that require manual feature engineering. Importantly, the researchers' model performed accurately in both training and testing, indicating it's learning clinically relevant patterns from the data, not subject-specific information.

Next, the researchers want to monitor how various treatments—such as surgery and vocal therapy—impact vocal behavior. If patients' behaviors move from abnormal to normal over time, they're most likely improving. They also hope to use a similar technique on electrocardiogram data, which is used to track muscular functions of the heart.

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

Provided by Massachusetts Institute of Technology

Citation: Automating artificial intelligence for medical decision-making (2019, August 6) retrieved 19 April 2024 from <https://techxplore.com/news/2019-08-automating-artificial-intelligence-medical-decision-making.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.