iHealthBeat
Reporting Technology's Impact on Health Care

by Ray Guzman
Friday, October 16, 2015

Seizures and Mosquitoes: The Rewards of Working Smarter With Data Science

Big Data has been defined as the oil of the 21st century. Just as crude oil must be refined into gasoline to power our cars, the large, complex data sets big data comprise aren't much use until they're honed into actionable insights. Data science deploys a range of tools—from crowdsourcing to visualization -- to capitalize on the promise of big data.

Until recently, the U.S. health care system has been less than immersed in the big data revolution. What drives most providers are issues of compliance or the avoidance of financial penalties, such as those associated with readmission.

But the landscape is changing. Instead of looking at a population of diabetics and asking retrospective questions, such as "What percentage required hospitalization? For how long?," big data can help answer specific questions around a given population, diagnosis or risk factor. This intelligence is the fuel that providers need to produce the most effective, personalized intervention.

The elegance of data science provides the kind of granular detail needed to answer very specific questions, but it can also be harnessed to tackle some of the world's thorniest health challenges.

Two case studies illustrate how data science is used to gain insight into world health issues.

Crowdsourcing To Predict Epileptic Seizures

The Mayo Clinic threw out a problem to a group of data scientists across the globe. Together with the American Epilepsy Society, it used the wisdom of the crowd to develop computer algorithms for the detection, prediction and prevention of epileptic seizures.

The particular needle-in-a-haystack problem was this: Physicians were looking to apply DBS (deep brain stimulation) as a treatment for epilepsy, and they needed highly specific information on the data surrounding pre-seizure brain activity.

To start the project, EEG data from both human and canine subjects were provided to approved teams. The goal was to find a variable to identify an algorithm to predict who will have a seizure and who will not.

Two hundred teams took up the 60-day challenge. The benchmark for the project was 65% accuracy, and the top two teams hit 95%.

The algorithm development process, as designed by one team, involved creating an artificial neural network that acts as a starting point for analysis. The team also used what it had learned from a previous challenge -- locate black holes based on the halo effect of starlight by separating signal from noise. The team has documented the algorithm for Benjamin Brinkmann of the Mayo Systems Electrophysiology Lab so Mayo researchers can use it for further innovation in epilepsy treatment.

But Mayo isn't keeping all this information to itself. Perhaps the greatest promise this and other data science competitions hold is that all of the data and algorithms will be shared openly online. Crowdsourcing creates a new approach for medical research that holds promise for the one in 26 people who suffers from epilepsy worldwide, and the one-third of those individuals whose seizures are not sufficiently treated with medication or other therapies.

Predicting Outbreaks of West Nile Virus

Here's another world health issue that big data addresses. West Nile virus is most commonly spread to humans through infected mosquitoes. From the time it was first reported in the U.S. in 1999 through 2012, treating the West Nile virus cost $778 million in expenses and lost productivity. Twenty percent of people who become infected with the virus develop symptoms ranging from a persistent fever to serious neurological complications that can result in death.

City public health systems have established comprehensive surveillance and control programs for West Nile since roughly 2004, but very few have developed a process to predict where the virus might occur.

Again, mountains of data -- including weather, location, testing and spraying -- are corralled to develop an algorithm to predict locations where West Nile is likely to become an issue. The algorithm predicts when and where different species of mosquitoes will likely test positive for West Nile virus. Families can protect themselves more aggressively and providers can be better prepared.

These findings will be presented at a global health conference in Sri Lanka later this year.

Refining big data into information we can use has enormous implications. For data scientists tackling health care issues, the opportunities are as big as curing epilepsy and as widespread as the summer's mosquitoes.

What is most important is finding ways to affordably and rapidly get answers to questions that allow health care to become a proactive industry rather than reactive.

When health care organizations of all sizes are faced with a rising number of challenges, better data to support clinical, financial and operational challenges are a requirement. Providing access to "practical data science" is an important component to ensure the big data tide, combined with data science, lifts quality care delivery and operational performance.