Prognos Looks to Make Sense of Unstructured Lab Data for Disease Diagnosis

Jul 10, 2017 | Neil Versel

With the recent hirings of three new executives, Prognos, a company that applies artificial intelligence to clinical diagnostics and laboratory testing, is eying what it called a "rapid growth stage."

Two weeks ago, the firm, which is based in New York, said that it had hired Lisa Kerber as chief operating officer, Stephen Silvestro as chief commercial officer, and Fernando Schwartz as chief data scientist. Schwartz, in particular, has some big plans for the company.

"Our vision for deploying AI in our dataset is to beat the MD at predicting disease. We don't want to replace the MD. We want to come alongside and empower the MD to have a bird's-eye view," he said.

In its seven years in business, Prognos has built a registry of clinical diagnostics and lab information that now encompasses more than 10 billion records on 160 million patients. Data comes from customers across the payer, life sciences, and medical diagnostics industries.

"No one person can possibly handle this type of data. That's why we need artificial intelligence," Schwartz said. "We can't achieve our goal without it because no one can digest the 10 billion records."

In turning to computers for help, Schwartz channeled the recently deceased Lawrence Weed, the medical informatics pioneer who envisioned clinical decision support 60 years ago. "The unaided mind does not know what data to collect, and does not see many of the significant relationships buried in whatever data are collected," Weed said in a 2004 interview.

Schwartz said Prognos is simply addressing the needs of its customers by ramping up its AI. Actually, AI is a fairly new offering for the privately held company, which is mostly funded by a $23 million investment from venture capital firm Safeguard Scientifics and the Merck Global Health Innovation Fund.

Artificial intelligence was not in the original business plan, according to Prognos Chief Marketing Officer Anne Bentley. "It came along as the technology improved. It came along as the computational power improved," she said.

Prognos was founded in 2010 by CEO Sundeep Bhan and his cousin Jason Bhan, a family physician who serves as chief medical officer.

Sundeep Bhan is, in Bentley's words, a "serial entrepreneur." He sold his previous company, Medsite, to WebMD in 2006 for $41 million; Bentley also was at WebMD at the time. The Bhans came into Prognos concentrating on clinical lab data only.

According to Bentley, Prognos was a data-driven company at the start, while its use of AI really began a few years ago. "Until you build up a mass of data, you really can't do much with AI. We needed to build up our volume to actually start to apply artificial intelligence technology," she said.

At Prognos, Bentley explained, "We started working with clinical lab data, clinical lab results, understanding what's important in terms of the patient journey from the time of symptoms appearing, understanding where the value might be."

About two years ago, as Prognos brought on more life sciences companies as clients, a new opportunity arose. "We actually identified that there might be similar business needs to have a better understanding of what's going on with clinical lab results and clinical data among payers, so we went into that market," Bentley said.

Bentley formerly worked for a pharmacy benefits manager that processed billions of claims through its mail-order pharmacy and retail network. "In that business, the claims structure, the data structure, is all standard." Claims data is coded in ICD-9 or ICD-10, the standard coding system for medical billing in the US; ICD-10 superseded version 9 in 2015.

"There's nothing like that in the clinical lab space," she said, at least as it applies to in vitro diagnostic devices. Lab results are coded with the Logical Observation Identifiers Names and Codes (LOINC) system. A newly published standard seeks to link IVD tests with LOINC codes, but that standard has not been around long enough to attract much adoption.

Uncoded data is "error-prone" and nonstandard, Bentley noted.

"We started adopting, as the technology came along, artificial intelligence technology in order to make sense of the data itself," she said. "The easiest possible application of AI is to use simple data at a large volume. When the data gets complicated, then much of the power of AI gets involved in detangling that."

Indeed, many healthcare datasets are full of unstructured text. "There is a lot of data out there, but not all data is created equal," Schwartz said.

Although IVD outputs are nothing but numbers and thus easy to structure, there is plenty of variability, Schwartz said, in large part because of consolidation in the laboratory industry. Mergers typically bring together organizations with different data formats.

"Their technology on the back end is playing catch-up," Schwartz said of smaller labs. "They never really wanted necessarily to sell data. Their business is about the results." They just wanted to be able to communicate results to patients and healthcare professionals.

Because of the large variability in data input, "We spend a great deal of time cleaning up and harmonizing our data," Schwartz said.

Prognos applies machine learning in two ways, according to Schwartz. The first is to prepare data for consumption by another machine learning algorithm.

Secondly, the technology makes projections by applying machine learning to look for patterns and trends in the harmonized data for predictive analytics. "Our dataset has both a probabilistic angle to it and a deterministic angle to it," Schwartz explained.

Pointing to pancreatic cancer as an example, he said that diagnosis of the disease has changed little during the past three decades, and most patients diagnosed with it are already in the advanced stages of the illness.

"But that is based on the traditional way of assessing information, and we are working toward pushing this to the limits to see what is probabilistic," he noted. "If you can beat the doctor in that diagnosis, you can have a huge impact on a large population segment."

Currently, Prognos focuses on three areas: rare disease, oncology, and chronic diseases with comorbidities typically managed by primary care physicians, including diabetes.

"Because of the growth of our data, the application of artificial intelligence, our vision is really to push that predictive capability to the point where decisions can be made sooner. We have seen some of that today, where our technology is identifying patterns in populations where if treatment could be applied sooner, we'd actually change the course of the disease," Bentley said.