Analytics Week – Implementing Personalized, Precision Medicine with Artificial Intelligence and Semantic Graph Technology

Personalizing healthcare services for individuals creates several demands on data-driven functions in the medical field. Healthcare organizations are tasked with integrating structured, unstructured and semi-structured data, storing and cataloging them in relevant ways across use cases and locations, and leveraging emerging AI techniques for predictive capabilities which could potentially save lives.

Most of all, this process must occur in time to make a difference for patients.

According to ***Montefiore Health System*** System Senior Vice President and Chief Medical Officer Andrew Racine, who spoke at a ***recent event for the unveiling of Intel's Xeon Scalable Processors***, all of these measures must be implemented so providers can: "use information in real time to make clinical decisions that are going to allow us to intervene with patients and prevent them from having adverse outcomes."

Montefiore is currently engaged in such an undertaking with a ***Semantic Data Lake for Healthcare***(SDL). The SDL is powered by ***Franz's AllegroGraph***, architected by Intel, and fortified by Cloudera's Hadoop distribution. By merging a unique set of data management techniques with some of the most pertinent technologies across the data landscape, Montefiore is seeking to tailor its medical treatment and diagnoses for individual patients.

Streaming Patient Data

Presently, the most eminent manifestation of the SDL is a process in which patient data in Montefiore's Intensive Care Units (ICUs) are used to "predict who among those patients will [in] 24, 48, [and] 72 hours, from the current moment, be at risk for respiratory failure," Racine revealed. Those patients may be hospitalized for any variety of medical conditions. A chief component of Montefiore's ability to achieve its objective is the analysis of patients' "real-time streaming data", according to Racine. The processing power supporting this application of the SDL is considerably strengthened by Intel's Xeon Scalable Processors, which were expressly designed to deliver the scalable performance, consistent security and agility required of such critical decision-making. The incorporation of this hardware with the SDL's semantic technology is pivotal to providing trenchant data-management systems for time-sensitive decisions affecting patient health. These characteristics of the SDL are what Racine described as indicative of differentiators between past and future healthcare systems:

"From our standpoint, what's going to distinguish advanced health systems in the future from others is the realization on the part of those systems that what they are fundamentally is information management systems. They're going to need to have the ability to gather, to store, to aggregate and to analyze large amounts of information from data matrices on vast numbers of patients."

The Back End

Successfully determining which patients will be in danger of respiratory failure creates a number of demands on both technological and medical processes. Montefiore uses predictive analytics algorithms to determine which patients are in jeopardy of respiratory failure, some of which involves machine learning. Moreover, the care facility must contextualize the patient-generated data from ICUs with a host

of other data germane to individual patients. That data may involve sources pertaining to socio-economics, genetics, family history, previous hospitalizations, medical publications, and others apropos for the patients' condition. Informing streamed patient data with these sources presents a number of concerns on the backend, which are readily addressed with the underlying semantic graph database technology. The SDL for healthcare is aided by an ontological pipeline based on a number of standards (both industry ***and W3C based***) which model data in a uniform way—regardless of source, structure, or any other inherent characteristic. In this regard, the SDL is also substantially strengthened by a sophisticated system of classifications and taxonomies which provide a consistency in the terminology of the data involved, so there is a common meaning across such data. In a recent video, Montefiore director of Clinical Research Informatics and Albert Einstein College of Medicine Dr. Parsa Mirhaji stated the SDL has "semantic aspects, which is about making information more meaningful using computational algorithms. It borders with Artificial Intelligence." The semantic graph database can also accommodate the scale required of Montefiore's predictive big data analytics, which enables "us essentially to do these analytics much more quickly, at a lower cost," Racine noted.

Clinical Recommendations

At present, the SDL's predictive prowess is also used to formulate clinical recommendations for patients determined to be in danger of imminent respiratory failure. Practitioners can decide whether or not to use those recommendations. Each patient outcome is monitored and input into the SDL's machine learning stack so that it increases its knowledge with each subsequent success or failure. According to Racine, the future of the SDL for healthcare could very well involve chronic issue diagnosis and, perhaps, chronic care management: "The real challenge for us is not to be able to predict who's going to get into trouble 48 hours from now—two days from now. For us, the challenge is to be able to predict who's going to

develop chronic conditions two years from now.”