Walter Tan
CS410: Text Information Systems
Technology Review

Amazon Comprehend Medical

When we think of Big Data, we can look at how organizations can leverage large amounts of data to find ways to improve operational efficiency or how to gain more insights about their customers. As a result, organizations are able to cut operating costs or increase their revenues. What about for industries such as Healthcare? Within healthcare, hospitals, pharmaceutical companies, and clinics alike can utilize Big Data to make sense of patient information so that this could be used to create better treatment plans for patients and to detect medical conditions in its early stages. Before the analysis can be completed and conclusions can be drawn, there are a number of challenges hospitals and healthcare clinics face in the initial stages of data collection. “Every year hospitals produce over 1.2 billion documents containing unstructured information” (“An Introduction”). The amount of data is massive and to process this data can be very time consuming. Data by itself is not useful, however the usability of large amounts of data can be improved by creating context through finding relationships between the provided text or through data categorization. One way of accomplishing that is through the use of Amazon Comprehend Medical. This is an API that can transform unstructured data into indexed, searchable, and structured data which can be used to drive decisions and to allow for improved patient care and efficiency (“An Introduction”). In this review, I will go over how Amazon Comprehend Medical works, its use cases, and its advantages and disadvantages.

Unstructured data can be in the form of physician notes, prescription information, admission notes, lab tests, and patient health data (“What is Amazon”). This data can be fed into Amazon Comprehend Medical which uses Natural Language Processing to extract important and relevant information from these types of free-form medical text (“Amazon Comprehend Medical). Amazon Comprehend Medical can then analyze and categorize the information with predetermined medical labels to build relationships between text and create indexes. Data can be categorized in the following entities: medical conditions, treatment procedures, anatomical terms, and protected health information (“An Introduction to”). For example, if a physician meets with a patient and writes the following into his or her appointment notes, “John Doe Age 55 experienced right arm pain and has taken a 4000mg of Ibuprofen daily.” Amazon Comprehend Medical will then take this data and categorize it using predetermine medical labels and will store this information. Anatomy -> System Organ Site: Right Arm, Medication -> Ibuprofen, Dosage-> 4000mg, Frequency: Daily, Protected Health Information -> Name, Age: 55 (Friedman, Ravi, Rattan, 00:08:24 - 00:09:50) . The information provided is now grouped, indexed, and searchable making this information more useful and valuable. In addition, relationships are maintained while using Amazon Comprehend medical. In the previous example, we see that for the Ibuprofen this is categorized as Medication. Within

Ibuprofen, medication dosage and frequency are both tied to the medication Ibuprofen. So now this offers more levels of granularity through search ("An Introduction"), for example if you need to find patterns with all patients who have taken Ibuprofen the data can be filtered in order to find what you are looking for.

As mentioned earlier there a wide number of uses cases that can benefit and leverage the capabilities of Amazon Comprehend Medical. One use case is early detection of medical conditions which would help immensely with saving lives and improving overall patient treatment costs. The sooner a medical condition is detected (depending on the medical condition) the likelier the patient can survive and to be able treat the patient with necessary medication or other therapies. Now that the data is structured and categorized it makes it easier to model the data and to find relationships between entitles such as age, symptoms, anatomy info, laboratory test results, and diagnosis of past patients. With this newly created model an application can be created to help patients with early diagnosis or Amazon Comprehend Medical can be integrated to an existing system. If a patient was overweight and was diagnosed with high blood pressure, then they could be recommended to get screened for another medical condition such as diabetes. The doctor may diagnose them as prediabetic which they could take steps to control before it becomes a larger issue. In addition, since the data is indexed and searchable, it would make it easier for doctors to search for past patients with similar symptoms and compare their information and how the patient was treated.

Another use case is finding better treatment plans for patients. In order to accomplish this the following entities will need to be utilized, medical conditions, patient information, diagnosis, allergy information, and treatment. A model can be created using these entities and to identify relationships between these types of data. Relationships between the variables can be mapped and this could be used to see if there are positive correlations between what type of treatment plan was given to the patient and what the outcomes were. In addition, we can review existing data to find others with similar patient criteria to find treatments that have helped similar patients in the past.

Amazon Comprehend Medical provides a number of advantages which include serverless architecture which would not require for organizations and other personnel to maintain a database which may help reduce costs in the long run. This technology is flexible enough so it can be easily integrated with existing and new applications through API calls which allows for seamless experience. In addition, this technology allows medical organizations to quickly gather information about the patient and will help save a lot of time during the collection phase. There are also a few disadvantages of the technology that could arise. Similar to other software, there could be instances where it will not work as expected and, in this case, there could be a possibility of incorrectly categorizing data depending on the way the notes were interpreted in the physician notes. If left uncaught then, it could present an issue when creating models. The result is producing inaccurate models.

In summary, Amazon Comprehend Medical allows hospitals, medical professionals, and other organizations to quickly collect and gather data about patients, thereby helping create actionable decisions based on the data presented. In doing so these organizations can fully utilize the data to integrate with existing applications and to find insights about patients and treatments. This will ultimately improve overall patient care for the future.

Works Cited

“Amazon Comprehend Medical.” *AWS,* Amazon. https://aws.amazon.com/comprehend/medical. Accessed 12 November 2020.

“An Introduction to Processing Unstructured Medical Data using Amazon Comprehend Medical.” *Health IT Analytics,* Xtelligent Healthcare Media. https://d1.awsstatic.com/whitepapers/HITA%20-%20Amazon%20Comprehend%20Medical%20-%20AWS.pdf. Accessed 12 November 2020.

Freidman, Aaron. Ravi, Arjun. Ratan, Ujjwal. "An Introduction to Processing Unstructured Medical Data using Amazon Comprehend Medica." , uploaded by Xtelligent Healthcare Media, 02 May 2019, https://vimeo.com/333817612.

“What is Amazon Comprehend Medical?.” *AWS,* Amazon. https://docs.aws.amazon.com/comprehend/latest/dg/comprehend-med.html. Accessed 12 November 2020.