Treating cancer with the power of cloud

Cancer Commons is a nonprofit network of patients, physicians, and scientists dedicated to helping patients identify and access the best personalized treatment options. This write up describes how the Amazon Web Services (AWS) compute infrastructure, combined with services like Amazon Comprehend Medical, enable physicians and patients to leverage the collective knowledge of the world's top institutions.

Metastatic cancer is one of the most devastating experiences one can face. People with cancer may be overwhelmed and confused by their treatment options. Often, they've already tried several treatments. Cancer Commons helps patients and caregivers navigate a cancer diagnosis by helping them understand their disease and informing them about additional testing and treatment options available. The options are curated by expert scientists and oncologists and refined based on actual clinical results. When necessary, we convene virtual tumor boards (VTBs), where nationally recognized experts in a patient's specific cancer will refine the options based on an individual's medical history and preferences. We also facilitate access to the recommended tests and treatments through clinical trials, expanded access, or reimbursement support.

Cloud computing turns data into impact and research into life-saving treatments

Every person's cancer is unique and requires an individualized approach to treatment. The best treatments involve intelligent combinations of therapies, and there are far too many plausible regimens to test in clinical trials. Cancer Commons and our technology partner xCures are solving this problem by integrating clinical research and care to continuously learn from all patients, on all treatments, all the time. This involves a lot of computing and data.

Services used in this use case

The patient data that flows into the system from electronic health records uses natural language processing (NLP) that relies in part on **Amazon Comprehend Medical.**

Throughout the oncology community, physicians reach out to colleagues to seek their opinions about how to treat a patient. This can be done informally via phone, email, or text, or in a more organized fashion, such as our VTBs. Unlike traditional tumor boards, where physicians at one institution meet in person to discuss their challenging cases, VTBs include experts from multiple institutions and multiple disciplines. They meet asynchronously, using a Slack-like platform to track simultaneous discussion threads covering multiple cases and treatment options. The group's collective reasoning about treatment options and rationales is captured and used to refine our options library. In this way, it can inform the treatment of other similar patients.

VTB system is developed on AWS, using a HIPAA-compliant architecture. We also rely on **Amazon Relational Database Service (Amazon RDS)** for the underlying PostgreSQL database services, so we don't need to worry about backup and replication. This process can be moved to **Amazon Neptune** in order to scale to take advantage of its persistent, performant graph database.

Bayesian models are trained on treatment rationales and outcome data, to predict the safety and efficacy of a given treatment on a patient or cohorts. All of the cancer treatment modeling runs on AWS. The discussion that happens in the VTBs is mined, using natural language parsing and other artificial intelligence (AI) techniques, and a recommended treatment plan is created from that information. A Bayesian hierarchical model of disease progression is created, where the structure is driven by the terms from the VTBs. The model is trained using data that comes from both the external medical literature and the treatment histories of those patients who have gone through VTBs and who have generated follow-up data. Training these models is a compute-intensive task. While we currently do this on standard AWS instances, we expect to fine-tune this using GPU clusters and tools like **Amazon SageMaker**.
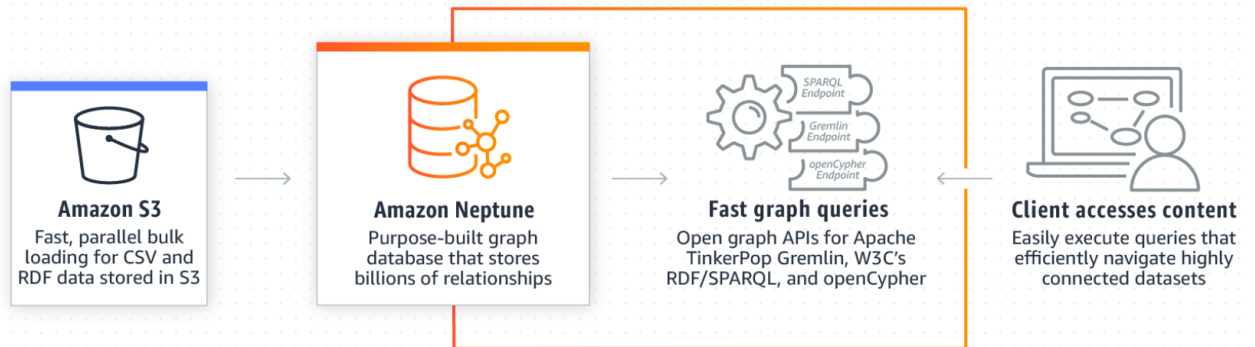
Overview of Services :

- Amazon Comprehend Medical -

- Amazon RDS -

- Amazon Neptune - Amazon Neptune is a fast, reliable, fully managed graph database service that makes it easy to build and run applications.

Amazon S3
Fast, parallel bulk loading for CSV and RDF data stored in S3

Amazon Neptune
Purpose-built graph database that stores billions of relationships

Fast graph queries
Open graph APIs for Apache TinkerPop Gremlin, W3C's RDF/SPARQL, and openCypher

SPARQL Endpoint
Gremlin Endpoint
openCypher Endpoint

Client accesses content
Easily execute queries that efficiently navigate highly connected datasets

- Amazon EC2 Instances - Amazon Elastic Compute Cloud (Amazon EC2) is a web-based service that allows businesses to run application programs in the AWS Public Cloud. Amazon EC2 allows a developer to spin up virtual machines, which provide compute capacity for IT projects and cloud workloads that run with global AWS data centers.

- Amazon Sagemaker - Amazon SageMaker is a fully managed machine learning service. With SageMaker, data scientists and developers can quickly and easily build and train machine learning models, and then directly deploy them into a production-ready hosted environment. It provides an integrated Jupyter authoring notebook instance for easy access to your data.

References :
https://aws.amazon.com/blogs/publicsector/treating-cancer-power-cloud/

Submitted By:
Amogh Garg - 2020UCO1688
Arushee Verma - 2020UCO1695
Nikhil Kumar - 2020UCO1708
Shubham Sharma - 2020UCO1705