Kristopher Gallagher - CS410 - Technology Review

Most applications today rely on cloud providers to develop and deliver their applications. All of the major providers are also offering Artifillience Intelligence(AI) tools for data scientists and application developers, including tools for text data retrieval and analysis. Since most development is done using services of cloud providers it is useful to analyze the technology available through the top providers. Using their tools can be efficient for developers and data scientists since most of their data is already stored and managed by these platforms. However it is very difficult to come through the documentation to understand what provider to pick. This paper will analyze Amazon Web Services (AWS) and Google Cloud Platform (GCP). These platforms were chosen because GCP is widely regarded as having the best AI offering.(Jones) AWS is also strong in this area and the overall market leader in cloud technology. (Jones) This paper will first provide a summary of tools provided by AWS, then tools provided by GCP and then a comparison of the two. The intent is to help summarize the software toolkits offered by both platforms to help inform a decision on which provider to choose when starting out on a project.

AWS has two major categories of machine learning tools. The first are platforms that allow the machine learning engineers and data scientists the tools they need to manage and deploy machine learning models. These platforms abstract the common elements it requires to make all machine learning models and allow the developer to focus just on the model development. Additionally it allows easy integration into existing applications running on AWS to allow end users to quickly benefit from the model developed. The major AWS machine learning platform is Sagemaker. This handles a variety of machine learning tasks including text analysis. Most common algorithms are already built-in, for example BlazingText is the AWS implementation of the Word2Vec algorithm. In addition to Text classification AWS has built in algorithms for classification (KNN), Blazing Text, Topic Modeling (LDA (Linear discriminant analysis), NTM(neural Turing Machine)), Clustering (KNN(K-nearest neighbor), K-Means), and Feature Reduction (PCA (principal component analysis, Object2Vec) (Amazon Web Services). A safe summarization is that AWS takes the state of the art models available and builds the algorithms into its sagemaker platform. One piece of the infrastructure that AWS provides an IDE optimized for their sagemaker platform.

The second category of tools that AWS provides is pre-trained models. This category of tools are meant for developers that do not have the expertise, time or resources to train their own model. While this takes a lot of the heavy lifting of training your model off of the developer's shoulders it comes at the cost of the flexibility to train a model for your specific use case. These tools are typically implemented by the developer sending over data to an HTTP end-point. That data will be analyzed by a generic AWS model and return the results of that model. There are several text based pre-trained models that AWS has for developers. There are tools for speech-to-text, translation, and for chatbot interactions. The main tool to conduct text analysis is Amazon comprehend. Amazon comprehend covers many of the analysis methods described in this course including sentiment analysis, topic modeling and key phrase detection (Amazon Web

Services). Taking a deeper look at the sentiment analysis it uses an LDA learning model and requires the developer to upload at least 1,000 documents to train the model along with an object that contains the list of topics for the documents (Amazon Web Services). Additionally AWS provides industry specific models for certain segments that have very unique language like medicine.

Finally, it is worthwhile to mention that AWS has a marketplace to allow developers to shop for pre-trained models that meet their specific needs. The vendors of these models are required to implement them via the above referenced Sagemaker platform so it easily integrates into other AWS environments. This provides the developer the option to look for more specific pre-trained models that can be used for their specific application but are too niche for Amazon to develop internally.

Next we will analyze the tools provided by Google Cloud Platform. Again Google works under the same principle of providing two categories of products. The first being a platform for developers to build, train and deploy their own models and the second being the pretrained models developers can use to integrate into their existing software. Again Google touts the ability to easily integrate (Google Cloud Platform) into already deployed applications on GCP and being able to integrate the ML model into an end-to-end workflow. GCP's cloud machine learning platform is AutoML and VertexAI. (Google Cloud Platform) GCP's VertexAI provides significantly less built-in algorithms and no algorithms specifically built for text analysis. Although these algorithms are not built-in, a developer would be able to import these libraries but it would require significant manual work. Google Cloud's also has a few pretrained models to support NLP functions. Most of their pretrained models are around pre-treating the data prior to being put into an ML pipeline. For example they have syntax analysis and parsing pre-trained models but little once the data is ready for analysis. Their platform is designed to then move the developer over to their coure AI platform for development. This is targeted for a more sophisticated developer use case. They do have industry specific products. For example contact center sentiment analysis. Overall Google Cloud is much more targeted to a sophisticated set of developers who want to train their own models.

Next we will compare the two competitors on their performance, prices and product lineups. The first category of product lineups AWS has a full set of features and functionality. AWS has been in the cloud market much longer than GCP and it shows in the maturity of their offering. AWS provides many built-in algorithms that encompass the majority of the state of the art models. GCP only provides tools to help the developer clean their data and translate it into something usable in an algorithm but not much in terms of built in algorithms. For the more sophisticated developer this should be very suitable. AWS also provides a more pre-trained model where the developer can simply pass in data to an HTTP end-point and get a response from the analysis AWS will do from a more generic model. Again, this will be less of a concern for a more sophisticated developer. Performance has been hard to get detailed numbers on. Since this is highly confidential data, it is closely guarded by the companies. However, from various blog posts it seems the sentiment lies with GCP as being a faster platform. This can be attributed

to their custom built GPU's, their "TPU - Tensor Processing Units" (Jones). While many companies simply repurpose GCU's to meet their Machine learning needs. Google has designed their very own chip optimized for deep learning workloads. This has provided them a leg up in terms of performance against AWS that is using a third party chip. Finally on price Google has a slight advantage here. While again the prices are highly variable and depend on the specifics Google tends to come in cheaper than AWS (Jones). AWS and GCP are both very capable text analysis and NLP platforms but target different user bases. AWS has a large number of tools for developers who do not have the time or expertise to develop and tune sophisticated models on their own. However Google offers better performance and price for the core feature set that will easily suit the needs of a developer that just needs a platform to handle the infrastructure elements of their project. Hopefully this technology review has been helpful in guiding developers toward picking a provider for their text analysis and retrieval needs.

Works Cited

Amazon Web Services. "Amazon Comprehend Developer Guide." *Amazon Comprehend*

Developer Guid, 2021,

https://docs.aws.amazon.com/comprehend/latest/dg/comprehend-dg.pdf. Accessed 06

November 2021.

Amazon Web Services. "Amazon Sagemaker Developer Guide." *Amazon Sagemaker Developer*

Guide, 2021,

https://docs.aws.amazon.com/sagemaker/latest/dg/sagemaker-dg.pdf?icmpid=docs_sage

maker_lp. Accessed 7 November 2021.

Google Cloud Platform. "Introduction to AI Platform." *Introduction to AI Platform*, 2021,

https://cloud.google.com/ai-platform/docs/technical-overview. Accessed 07 Novebmer

2021.

Jones, Edward. "Google Cloud vs. AWS 2021." *Google Cloud vs. AWS 2021*, 2021,

https://kinsta.com/blog/google-cloud-vs-aws/. Accessed 7 November 2021.