

Shivang Saxena

CS 410

Technology Review

11/7/2021

Comparison of Amazon and Google's Natural Language Processing Cloud Offerings

The proliferation of the cloud has resulted in many computation-heavy tasks being offloaded from being run on local machines to cloud offerings, known as “SaaS” (Software-as-a-service). The SaaS market is over \$120 billion and is only growing (Dale, R). Cloud SaaS offerings free developers from having to set up and maintain complex infrastructures and decrease time to market and prototype development time. Cloud SaaS offerings from Google and Amazon are now starting to offer NLP (Natural Language Processing) features such as entity analysis, sentiment analysis, and topic classifications, which allow developers to build sophisticated software rapidly while having appealing economics as well as seamless scalability thanks to the benefits of the cloud (Dale, R). This paper will compare and contrast Google and Amazon's NLP features related to entity analysis, sentiment analysis, and topic classifications, from both a product perspective (how do their offerings differ) and a scalability perspective (which offering is easier and more economic to scale on the cloud).

Google offers a “Natural Language API” on their Google Cloud platform that provides entity analysis, sentiment analysis, and topic classifications. Google's entity analysis cloud offerings can identify entities within documents such as receipts, invoices, and contracts, and label them with contextual information such as media type, persons involved, and dates of documents (*Google Cloud Natural Language*). Google supports multiple languages such as English, Chinese, French, German, Italian, and more, and provides sample REST code that can be used to perform entity analysis in languages such as Go, Java, Nodejs, and Python. From a

cost and scaling perspective, entity analysis on Google's Cloud Platform is very fair, as up to 5,000 document requests are free for entity analysis, and 5,000 to 1M requests only cost \$1. Each request is considered a document with up to 1,000 unicode characters. Google Cloud Platform also provides sentiment analysis features, which allows developers to easily understand the sentiment of documents, from feeling, attitude, and overall opinions expressed in documents. Similar to entity analysis, this is available in a variety of programming and spoken languages, and has identical price structure — free for 0 to 5,000 document requests, and \$1 for 5,000 to 1M requests. Finally, Google Cloud offers topic classification features. This allows developers to easily classify documents into 700 predefined categories. One limitation in this offering is that these topics are predefined and cannot be customized beyond what Google offers. Pricing and scalability options are slightly different from Google's other NLP offerings, as it's free for 0 to 30,000 requests, but then \$2 for 30,000 to 250,000 requests. Overall, Google's Cloud NLP offerings are vast and a combination of entity analysis, sentiment analysis, and topic classification should give developers lots of features to combine to create robust products with NLP. As discussed, Google offers generous free pricing tiers so that developers can start without friction and move onto paid tiers as necessary, which will be easy to scale in the cloud (*Google Cloud Natural Language*).

Amazon's NLP offerings are similar but differ from Google's in the way they are consumed, priced, and how they function. Similar to Google's Cloud Platform, entity analysis allows developers to provide documents and get entities such as organization, date, and location (*Amazon Comprehend Features*). Amazon cloud also offers a "confidence score" for each entity determined, which developers can use to set expectations. Unlike Google Cloud Platform, however, Amazon does not provide a free tier for entity analysis and has a flat price of \$0.0001 per "unit". Another difference in the pricing is a unit is considered 100 characters, compared to Google's Cloud's 1,000 characters per unit. This means that Google's offering is cheaper to get started but as you scale Amazon might end up being more economical. Amazon

cloud also offers sentiment analysis, which allows developers to understand the overall sentiment of a document. Unlike Google's offering that provides sentiment, opinion, feel, and attitude, Amazon's sentiment analysis only offers an overall positive, negative, or neutral sentiment analysis, along with a confidence score. The pricing for the sentiment analysis is the same as entity analysis described previously, however, unlike Google's offering, amazon does not provide a free tier. Finally, Amazon cloud also offers a comparable topic classification cloud offering. This feature allows developers to group documents by topic, and provides a confidence score. This feature has a few differences from Google's offering. First, it is not limited to a set of 700 predefined categories. Second, this feature allows both documents and keywords to be grouped by topic. The pricing structure is also different, as it is \$1 per "job", with each job being limited to 100MB of data. Beyond that, it is \$0.004 per MB of document data that needs to be classified for topics (*Amazon Comprehend Features*).

Both Google Cloud Platform and Amazon Cloud provide extensive NLP features such as entity analysis, sentiment analysis, and topic classifications. These features enable developers to rapidly prototype NLP products without needing to develop and maintain complex infrastructure and processing pipelines, and benefit from cloud scalability and great economics (and even free tiers for smaller use-cases). The major differences between these offerings come down to pricing, as Google offers much more free tiers compared to Amazon, as well as feature customizability. In general, it seems Amazon's NLP offerings are more configurable (for example, topic classification is not limited to 700 pre-defined categories) so developers who need more customization may look to Amazon's offerings, but may find their lack of free pricing tiers restrictive for rapid prototyping. However, both cloud offerings seem very robust and a great fit for those who want to rapidly produce an NLP product and perhaps one day scale it to millions of users.

References

Amazon Comprehend Features. (2014). Amazon. Retrieved November 7, 2021, from <https://aws.amazon.com/comprehend/features/>.

Google Cloud Natural Language. (n.d.). Google. Retrieved November 7, 2021, from <https://cloud.google.com/natural-language#section-1>.

Dale, R. (2015). NLP meets the cloud. *NATURAL LANGUAGE ENGINEERING*, 21(4), 653–659.
<https://doi-org.proxy2.library.illinois.edu/10.1017/S1351324915000200>