Technology Review: Amazon Comprehend vs Google Cloud Natural Language

Chee Tey (cheeht2@illinois.edu)

November 15, 2020

Abstract

Amazon Comprehend and Google Cloud Natural Language are popular text processing APIs provided by Amazon and Google. They are accessible solutions developed to speed up and simplify natural language processing (NLP) without specific domain knowledge in NLP.

1 Introduction

Natural Language Processing (NLP) helps computers interpret human language. It is a study of computer science and linguistic. The use of NLP recently exploded with many people inadvertently using it without realizing, such as auto-captioning of a video on YouTube, Siri, spam filtering of emails, etc.

In the past, people need to create complex machine learning models to implement an NLP solution. However, with the increasing popularity of Cloud, many Cloud Providers are introducing pre-trained machine learning models to implement some of the most popular NLP features. This meant that any developer can implement NLP solutions in their application without actually knowing NLP - they can simply access the service via an API. In this paper, we will focus on two of the most popular NLP services: Amazon Comprehend and Google Cloud Natural Language.

2 Amazon Comprehend

Amazon Comprehend[1] is a natural language processing (NLP) service provided by Amazon Web Services. Like most NLP services, it provides API-based pre-trained machine learning models to discover insights from structured and unstructured text. The API outputs its results in JSON format, which can be easily consumed by an application.

Amazon Comprehend's pre-trained features include:

1. Keyphrase Extraction
2. Sentiment Analysis
3. Syntax Analysis
4. Entity Recognition
5. Language Detection
6. Topic Modeling

Amazon Comprehend's other features which support custom models include:

1. Custom Entities
2. Custom Classification

3 Google Cloud Natural Language

Google Cloud Natural Language[2] is a natural language processing (NLP) service provided by Google Cloud Platform. It provides two types of services: Natural Language API and AutoML Natural Language. Natural Language API provides API-based pre-trained machine learning models to discover insights from structured and unstructured text. On the other hand, AutoML Natural Language allows users to build and deploy their own machine learning models to analyze documents, categorizing, identify entities, etc.

The Natural Language API (pre-trained models) supports:

1. Syntax Analysis
2. Entity Analysis
3. Sentiment Analysis
4. Content Classification

The AutoML Natural Language (custom models) supports:

1. Custom Entity Extraction
2. Custom Sentiment Analysis
3. Custom Content Classification

4 Analysis of Features and their Differences

We will be comparing the feature of both services with the same text:

"Elon Musk, the CEO of the electric car company Tesla, has been called a genius, a visionary. But this year he got less noticed for his brilliance than his behavior, with stories about self-inflicted wounds: like capricious tweeting and public pot smoking. The 47-year-old billionaire has said 2018 has been excruciating, the most... painful year of my career. When he joined as co-founder of the company 14 years ago, he had no experience in the auto industry. He was guided by a dream: to build cars that don't harm the environment in an effort to save the planet. But this year Musk had to save his company. Tesla had been losing money for nearly its entire existence, its debt is in the billions, and it was bleeding cash. Everything was riding on its ability to mass produce its new sedan: the Model 3, an affordable so-called everyman car."[3]

Here's an analysis all the features:

1. Syntax Analysis
   Both platform provides pre-trained syntax analysis capability. An output of this analysis for a portion of the first sentence:

	Elon	Musk	,	the	CEO
Amazon Comprehend	Proper noun	Proper noun	Punctuation	Determiner	Proper noun
Google Cloud Natural Language	Noun	Noun	Punctuation	Determiner	Noun

Table 1.0 Table showing the syntax analysis results

   Verdict: Both platforms provide accurate syntax analysis

2. Entity Analysis
   Both platform provides pre-trained syntax analysis capability. An output of this analysis for a portion of the first sentence:

	Tesla	Elon Musk	2018	47-year-old	Model 3
Amazon Comprehend	Organization	Person	Date	Quantity	Commercial item

Table 1.1 Table showing the entity analysis results for Amazon Comprehend

	Tesla	Elon Musk	2018	47	Model
Google Cloud Natural Language	Organization	Person	Number	Number	Consumer Good

Table 1.2 Table showing the entity analysis results for Google Cloud Natural Language

Verdict: Amazon Comprehend is better at entity analysis than Google Cloud Natural Language. For example, Comprehend recognized that "Model 3" is the entity here and tagged it as a "Commercial Item", whereas Google Cloud Natural Language analyzed it as "Model" and tagged it as a generic "Consumer goods".

3. Sentiment Analysis

	Neutral	Positive	Negative	Mixed
Amazon Comprehend	0.29	0.06	0.64	0.00

Table 1.3 Table showing the sentiment analysis results for Amazon Comprehend

	Sentiment Score
Google Cloud Natural Language	-0.1

Table 1.4 Table showing the sentiment analysis results for Google Cloud Natural Language

Verdict: Google Cloud Natural Language provided a better analysis on the sentiment of the text as it determined the sentiment to be between mixed and slightly negative. Amazon Comprehend is tilting towards Negative.

4. Content Classification
Amazon Comprehend does not offer this feature.

Google Cloud Natural Language:

Category	Confidence
Autos Vehicles, Motor Vehicles (By Type), Hybrid Alternative Vehicles	0.96
Business Industrial	0.65

Table 1.5 Table showing the content classification results for Google Cloud Natural Language

Verdict: Google Cloud Natural Language provided an accurate categorization of the text. Amazon Comprehend does not offer this feature.

5. Keyphrase Extraction
Google Cloud Natural Language does not offer this feature.

Amazon Comprehend:

Key Phrases	Confidence
Elon Musk	0.99+
the CEO	0.99+
the electric car company Tesla	0.99+
a genius	0.99+

Table 1.6 Table showing the Keyphrase Extraction results for Amazon Comprehend

Verdict: Amazon Comprehend provided an accurate depiction of the keywords of the text. Google Cloud Natural Language does not offer this feature.

5 Comparison of Services Provided by Amazon Comprehend and Google Cloud Natural Language

	Amazon Comprehend	Google Cloud Natural Language
Syntax Analysis	Yes	Yes
Entity Analysis	Yes	Yes
Sentiment Analysis	Yes	Yes
Content Classification	No	Yes
Keyphrase Extraction	Yes	No

Table 1.7 Comparison of Services Provided by Amazon Comprehend and Google Cloud Natural Language

6 Conclusion

While we only explored the pre-trained model features of both services, you can see the advantages it brings to the table to be able to bring commercial-grade NLP capabilities to your application without employing machine learning scientists to develop a model. For more advanced developers, both services allow you to host your model on their platform using AutoML to build custom models to suit your needs. Amazon Comprehend and Google Cloud Natural Language are excellent tools to leverage when considering adding NLP solutions to your application.

References

[1] Amazon comprehend.

[2] Google cloud natural language.

[3] Lesley Stahl. Tesla ceo elon musk: The 60 minutes interview. *CBS News 60 Minutes*, Dec 2018.