

URL to Colab:

AWS: <https://colab.research.google.com/drive/1L5FA0UiZhWkDMrvTN1IYox7t2ETB6jV7>

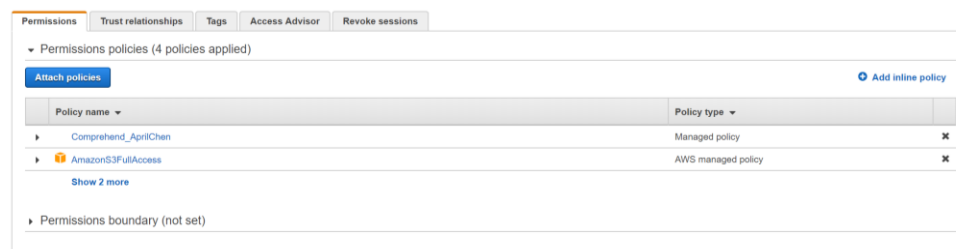
GCP: <https://colab.research.google.com/drive/1-waKbsVA0WOzdoVkk7X1mVgL3LMoH699>

Project Explanation:

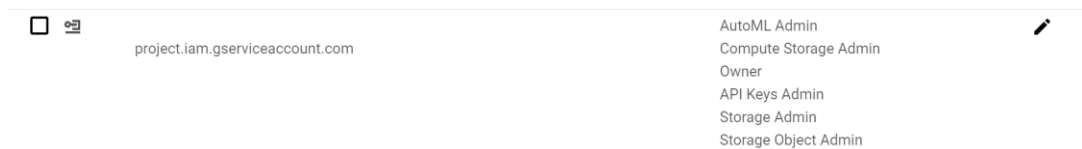
The project is to utilize AWS Comprehend and Google Cloud Platform Natural Language Processing to detect the customers' sentiments during the conversation with customer support in Twitter. By detecting the sentiments, we will have a better understanding of customers' attitudes and improve the quality of customer support in Twitter. The text conversation data is from Kaggle.

Access Google Cloud Platform (GCP) and Amazon Web Services (AWS)

Before accessing GCP and AWS, we have to grant full access to appropriate roles.



In AWS, we attach policies to the role to ensure that this role has the access to Amazon Comprehend AWS S3.



In GCP, we grant Admin access to API and Storage to the role we are going to use.

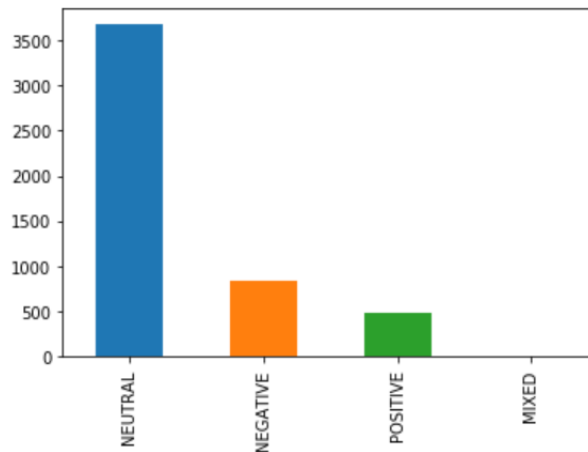
Data Ingestion from Google Cloud Storage and AWS S3

In the data ingestion process, we loaded the Twitter Customer Support Conversation csv file from AWS S3 and GCP Storage and converted it to pandas dataframe in Python. Then we filtered the whole data set by selecting only inbound texts, which are the responses from customers. Since there is a maximum limit times to call APIs of AWS and GCP, we only selected the first 5000 conversations in the whole dataset.

Project Conclusion:

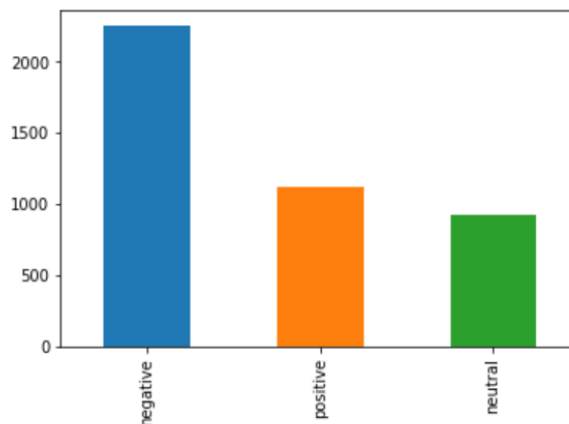
Result Analysis – AWS

From the below chart, which is the sentiment analysis result of Amazon Comprehend, we can see that almost 70% of the customers' sentiments among the first 5000 conversations are neutral, following by 20% of the customers' negative sentiments.



Amazon Comprehend Sentiment Analysis Result

Result Analysis – GCP



Google Cloud Platform Natural Language Processing Sentiment Analysis Result

Since GCP Sentiment Analysis responses only returns the sentiment scores, we defined the sentiment score that equals to 0 as neutral sentiment, and the score that is larger than 0 as positive sentiment; on contrast, smaller than 0 as negative sentiment. Surprisingly, we got totally different result from Amazon Comprehend, most customers' sentiments are negative, which take over 70% of the first 5000 conversations.

Recommendations to Management:

Since up to 70% of the first 5000 conversations from customers' sentiment scores are below zero, which indicate negative sentiments, we need to understand the causes of their negative emotions. It is also worthwhile to conduct analysis to see if there is any improvement on customers' sentiments as the conversation goes on with customer supports.

Since AWS and GCP have the limit times of free API calls, the limit in our analysis is that we can only focus on the first 5000 conversations. It is also worthwhile to improve our analysis by reviewing the whole conversation dataset to see if there is any bias in the first 5000 samples.