**Semantic Textual Similarity**

**Problem Statement**

Find similarity score between pairs of sentence. The score should be real-valued between 0 and 1.

**Approach**

BERT based approach was used to find the similarity score between pairs of sentence. “*distilbert-base-nli-stsb-mean-tokens*” was used to get the sentence embeddings, and cosine similarity was used to find the similarity score. The score with the corresponding Unique\_ID was written to a .csv file

**Programming Language :** python (Jupyter Notebook)

**Libraries :** sentence\_transformer, pandas, numpy

**Dataset :** Text\_Similarity\_Dataset.csv

**Environment :** Google Colab using CPU