

```
# importing necessary libraries
import pandas as pd
import numpy as np

# loading data set
df = pd.read_csv(r'/content/drive/MyDrive/Precily task/Precily_Text_Similarity.csv')

df.head()
```

	text1	text2
0	broadband challenges tv viewing the number of ...	gardener wins double in glasgow britain s jaso...
1	rap boss arrested over drug find rap mogul mar...	amnesty chief laments war failure the lack of ...
2	player burn-out worries robinson england coach...	hanks greeted at wintry premiere hollywood sta...
3	hearts of oak 3-2 cotonsport hearts of oak set...	redford s vision of sundance despite sporting ...
4	sir paul rocks super bowl crowds sir paul mcca...	mauresmo opens with victory in la amelie maure...

```
pip install transformers
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: transformers in /usr/local/lib/python3.7/dist-packages (4.24.0)
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (from transformers) (2.23.0)
Requirement already satisfied: importlib-metadata in /usr/local/lib/python3.7/dist-packages (from transformers) (4.13.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.7/dist-packages (from transformers) (6.0)
Requirement already satisfied: tokenizers!=0.11.3,<0.14,>=0.11.1 in /usr/local/lib/python3.7/dist-packages (from transformers) (0.13.0)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.7/dist-packages (from transformers) (1.21.6)
Requirement already satisfied: huggingface-hub<1.0,>=0.10.0 in /usr/local/lib/python3.7/dist-packages (from transformers) (0.11.0)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.7/dist-packages (from transformers) (2022.6.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.7/dist-packages (from transformers) (21.3)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.7/dist-packages (from transformers) (4.64.1)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-packages (from transformers) (3.8.0)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.7/dist-packages (from transformers) (4.5.0)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /usr/local/lib/python3.7/dist-packages (from packaging>=20.0) (3.0.9)
```

```
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata->transformers)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests->transformers)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests->transformers)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests->transformers) (2.9.0)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests->transformers) (3.0.4)
```

pip install sentence-transformers

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: sentence-transformers in /usr/local/lib/python3.7/dist-packages (2.2.2)
Requirement already satisfied: transformers<5.0.0,>=4.6.0 in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (4.11.0)
Requirement already satisfied: scipy in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (1.7.3)
Requirement already satisfied: torchvision in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (0.11.0)
Requirement already satisfied: torch>=1.6.0 in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (1.10.1)
Requirement already satisfied: nltk in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (3.7)
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (1.21.6)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (1.0.2)
Requirement already satisfied: sentencepiece in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (0.1.9)
Requirement already satisfied: huggingface-hub>=0.4.0 in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (0.11.1)
Requirement already satisfied: tqdm in /usr/local/lib/python3.7/dist-packages (from sentence-transformers) (4.64.1)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-packages (from huggingface-hub>=0.4.0->sentence-transformers) (3.4.0)
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (from huggingface-hub>=0.4.0->sentence-transformers) (2.28.1)
Requirement already satisfied: importlib-metadata in /usr/local/lib/python3.7/dist-packages (from huggingface-hub>=0.4.0->sentence-transformers) (4.2.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.7/dist-packages (from huggingface-hub>=0.4.0->sentence-transformers) (6.0.1)
Requirement already satisfied: packaging>=20.9 in /usr/local/lib/python3.7/dist-packages (from huggingface-hub>=0.4.0->sentence-transformers) (21.3)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.7/dist-packages (from huggingface-hub>=0.4.0->sentence-transformers) (4.4.0)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /usr/local/lib/python3.7/dist-packages (from packaging>=20.9->sentence-transformers) (3.0.9)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.7/dist-packages (from transformers<5.0.0,>=4.6.0->sentence-transformers) (2022.10.31)
Requirement already satisfied: tokenizers!=0.11.3,<0.14,>=0.11.1 in /usr/local/lib/python3.7/dist-packages (from transformers<5.0.0,>=4.6.0->sentence-transformers) (0.13.3)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata->huggingface-hub) (3.6.0)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages (from nltk->sentence-transformers) (1.2.0)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages (from nltk->sentence-transformers) (7.1.2)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests->huggingface-hub) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests->huggingface-hub) (2.9.0)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests->huggingface-hub) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests->huggingface-hub) (2022.9.24)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.7/dist-packages (from scikit-learn->sentence-transformers) (3.1.0)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in /usr/local/lib/python3.7/dist-packages (from torchvision->sentence-transformers) (9.0.1)
```

```

from sentence_transformers import SentenceTransformer, util          # I imported SentenceTransformer for loading-pre trained

model = SentenceTransformer('stsb-mpnet-base-v2')

sentence1 = "currently i am studying because tomorrow is my exam"
sentence2 = "I want to study as my exam will start soon"

# encode sentences to get their embeddings
embedding1 = model.encode(sentence1, convert_to_tensor=True)
embedding2 = model.encode(sentence2, convert_to_tensor=True)

# compute similarity scores of two embeddings
cosine_scores = util.pytorch_cos_sim(embedding1, embedding2)
print("Sentence 1:", sentence1)
print("Sentence 2:", sentence2)
print("Similarity score:", cosine_scores.item())

📄 Sentence 1: currently i am studying because tomorrow is my exam
Sentence 2: I want to study as my exam will start soon
Similarity score: 0.8182871341705322

sentences1 = df.iloc[:15,0].values      # as we have so much ammount of rows that consumes so much time so for timing i am tal
sentences2 = df.iloc[:15,-1].values
# encode list of sentences to get their embeddings
embedding1 = model.encode(sentences1, convert_to_tensor=True)
embedding2 = model.encode(sentences2, convert_to_tensor=True)
# compute similarity scores of two embeddings
cosine_scores = util.pytorch_cos_sim(embedding1, embedding2)

for i in range(len(sentences1)):
    for j in range(len(sentences2)):
        print("Sentence 1:", sentences1[i])
        print("Sentence 2:", sentences2[j])
        print("Similarity Score:", cosine_scores[i][j].item())

```

```
print()
```

Sentence 2: mauresmo opens with victory in la amelie mauresmo and maria sharapova won their opening matches at the t
Similarity Score: 0.22389613091945648

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: fosters buys stake in winemaker australia brewer fosters has bought a large stake in australia winemak
Similarity Score: 0.14864155650138855

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: jp morgan admits us slavery links thousands of slaves were accepted as collateral for loans by two banks
Similarity Score: 0.043976347893476486

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: howl helps boost japan s cinemas japan s box office received a 3.8% boost last year with ticket sales w
Similarity Score: 0.14737731218338013

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: ocean s twelve raids box office ocean s twelve the crime caper sequel starring george clooney brad pit
Similarity Score: 0.20683325827121735

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: uk broadband gets speed injection broadband s rapid rise continues apace as speeds gear up a notch. an
Similarity Score: 0.044648922979831696

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: protect whistleblowers tuc says the government should change the law to give more protection to employe
Similarity Score: 0.0958925411105156

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: lennon brands rangers favourites celtic s neil lennon admits rangers could be considered slight favouri
Similarity Score: 0.13561426103115082

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: unilever shake up as profit slips anglo-dutch consumer goods giant unilever is to merge its two manageme
Similarity Score: 0.15507932007312775

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: games firms face tough future uk video game firms face a testing time as they prepare for the next rou
Similarity Score: 0.038286689668893814

Sentence 1: sir paul rocks super bowl crowds sir paul mccartney wowed fans with a live mini-concert at american foot
Sentence 2: cactus diet deal for phytopharm a slimming aid made from a southern african cactus is set to be develope

Similarity Score: 0.002227015793323517

Sentence 1: india to deport bollywood actress india has ordered the deportation of iranian-born model and actress ne
Sentence 2: gardener wins double in glasgow britain s jason gardener enjoyed a double 60m success in glasgow in his
Similarity Score: 0.0016312227817252278

Sentence 1: india to deport bollywood actress india has ordered the deportation of iranian-born model and actress ne
Sentence 2: amnesty chief laments war failure the lack of public outrage about the war on terror is a powerful indic
Similarity Score: 0.28601133823394775

Sentence 1: india to deport bollywood actress india has ordered the deportation of iranian-born model and actress ne
Sentence 2: hanks greeted at wintry premiere hollywood star tom hanks was in london s leicester square for the uk pr
Similarity Score: 0.26502758264541626

Sentence 1: india to deport bollywood actress india has ordered the deportation of iranian-born model and actress ne
Sentence 2: redford s vision of sundance despite sporting a corduroy cap pulled low over his face plus a pair of dar

```
sentences1 = df.iloc[:15,0].values      # as we have so much ammount of rows that consumes so much time so for timing i am tal
sentences2 = df.iloc[:15,-1].values
# encode list of sentences to get their embeddings
embedding1 = model.encode(sentences1, convert_to_tensor=True)
embedding2 = model.encode(sentences2, convert_to_tensor=True)
# compute similarity scores of two embeddings
cosine_scores = util.pytorch_cos_sim(embedding1, embedding2)
```

```
score = []
for i in range(len(sentences1)):
    for j in range(len(sentences2)):
        print("Sentence 1:", sentences1[i])
        print("Sentence 2:", sentences2[j])
        print("Similarity Score:", cosine_scores[i][j].item())
        score.append(cosine_scores[i][j].item())
    print()
```

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: amnesty chief laments war failure the lack of public outrage about the war on terror is a powerful indic
Similarity Score: 0.17293456196784973

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: hanks greeted at wintry premiere hollywood star tom hanks was in london s leicester square for the uk pr

Similarity Score: 0.0814230740070343

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: redford s vision of sundance despite sporting a corduroy cap pulled low over his face plus a pair of dar
Similarity Score: 0.17443549633026123

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: mauresmo opens with victory in la amelie mauresmo and maria sharapova won their opening matches at the t
Similarity Score: 0.15619458258152008

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: fosters buys stake in winemaker australian brewer fosters has bought a large stake in australian winemak
Similarity Score: 0.3292539119720459

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: jp morgan admits us slavery links thousands of slaves were accepted as collateral for loans by two banks
Similarity Score: 0.034051235765218735

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: howl helps boost japan s cinemas japan s box office received a 3.8% boost last year with ticket sales w
Similarity Score: 0.05326668173074722

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: ocean s twelve raids box office ocean s twelve the crime caper sequel starring george clooney brad pit
Similarity Score: 0.06739998608827591

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: uk broadband gets speed injection broadband s rapid rise continues apace as speeds gear up a notch. an
Similarity Score: 0.10588280111551285

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: protect whistleblowers tuc says the government should change the law to give more protection to employe
Similarity Score: 0.17684036493301392

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: lennon brands rangers favourites celtic s neil lennon admits rangers could be considered slight favouri
Similarity Score: 0.20813332498073578

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: unilever shake up as profit slips anglo-dutch consumer goods giant unilever is to merge its two manageme
Similarity Score: 0.09808537364006042

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: games firms face tough future uk video game firms face a testing time as they prepare for the next rou
Similarity Score: 0.18071508407592773

Sentence 1: pearce keen on succeeding keegan joint assistant boss stuart pearce has admitted he would like to succee
Sentence 2: cactus diet deal for phytopharm a slimming aid made from a southern african cactus is set to be develope
Similarity Score: 0.17474573850631714

```
len(score)
```

```
225
```

```
print(score)
```

```
[0.06493174284696579, 0.05634061247110367, 0.12425386905670166, 0.09151121973991394, 0.07085824012756348, 0.17471799254
```

[Colab paid products](#) - [Cancel contracts here](#)

✓ 0s completed at 12:06 AM

