

model_03a_bert_base_triplet_margin_10

May 4, 2023

1 Model 03a

Evidence retrieval using a Siamese BERT classification model. This is similar to Model 01, however, it only uses official pre-trained models from hugging face.

Ref: - [Hugging face pre-trained models](#) - [Hugging face guide to fine-tuning](#) - [Hugging face guide to fine-tuning easy](#) - [SO Guide](#)

1.1 Setup

1.1.1 Working Directory

```
[ ]: # Change the working directory to project root
from pathlib import Path
import os
ROOT_DIR = Path.cwd()
while not ROOT_DIR.joinpath("src").exists():
    ROOT_DIR = ROOT_DIR.parent
os.chdir(ROOT_DIR)
```

1.1.2 File paths

```
[ ]: MODEL_PATH = ROOT_DIR.joinpath("./result/models/*")
DATA_PATH = ROOT_DIR.joinpath("./data/*")
NER_PATH = ROOT_DIR.joinpath("./result/ner/*")
```

1.1.3 Dependencies

```
[ ]: # Imports and dependencies
import torch
from torch.utils.data import Dataset, DataLoader
from torch.nn import Module, TripletMarginWithDistanceLoss
from transformers import BertModel, BertTokenizer
from torch.optim import Adam
from torch.optim.lr_scheduler import LinearLR
from torcheval.metrics import BinaryAccuracy, BinaryF1Score

from src.torch_utils import get_torch_device
```

```

import json
from dataclasses import dataclass
from typing import List, Union, Tuple
from tqdm import tqdm
import random
import numpy as np
from datetime import datetime
from math import exp

TORCH_DEVICE = get_torch_device()

```

/opt/homebrew/Caskroom/miniconda/base/envs/comp90042_project/lib/python3.8/site-packages/tqdm/auto.py:21: TqdmWarning: IProgress not found. Please update jupyter and ipywidgets. See https://ipywidgets.readthedocs.io/en/stable/user_install.html

```

from .autonotebook import tqdm as notebook_tqdm

```

Torch device is 'mps'

1.2 Dataset

```

[ ]: @dataclass
class ClaimEvidencePair:
    claim_id:str
    evidence_id:str
    label:int = 0

```

```

[ ]: @dataclass
class ClaimEvidenceTriple:
    claim_id:str
    pos_evidence_id:str
    neg_evidence_id:str

```

```

[ ]: class SiameseEvalDataset(Dataset):
    def __init__(
        self,
        dev_claims_path:Path,
        evidence_path:Path,
        device = None,
        verbose:bool=True
    ) -> None:
        super(SiameseEvalDataset, self).__init__()
        self.verbose = verbose
        self.device = device

        # Load claims data from json
        with open(dev_claims_path, mode="r") as f:
            self.claims = (json.load(fp=f))

```

```

    # Load evidence library
    self.evidence = dict()
    with open(evidence_path, mode="r") as f:
        self.evidence.update(json.load(fp=f))

    # Get a list of all evidences within the dev set
    self.related_evidences = sorted({
        evidence_id
        for claim in self.claims.values()
        for evidence_id in claim["evidences"]
    })

    # Generate the data
    self.data = self.__generate_data()
    return

def __generate_data(self):
    data = []
    for claim_id, claim in tqdm(
        iterable=self.claims.items(),
        desc="claims",
        disable=not self.verbose
    ):
        evidence_ids = claim["evidences"]

        # Get the positives
        for evidence_id in evidence_ids:
            data.append(ClaimEvidencePair(
                claim_id=claim_id,
                evidence_id=evidence_id,
                label=1
            ))

        # Get some negatives
        n_neg = 0
        for rel_evidence_id in self.related_evidences:
            if n_neg >= 10:
                break
            if rel_evidence_id in evidence_ids:
                continue
            data.append(ClaimEvidencePair(
                claim_id=claim_id,
                evidence_id=rel_evidence_id,
                label=-1
            ))
            n_neg += 1

```

```

        return data

    def __len__(self):
        return len(self.data)

    def __getitem__(self, idx) -> Tuple[Union[str, torch.Tensor]]:
        # Fetch the required data rows
        data = self.data[idx]

        # Get the label
        label = torch.tensor(data.label, device=self.device)

        # Get text ids
        claim_id = data.claim_id
        evidence_id = data.evidence_id

        # Get text
        claim_text = self.claims[claim_id]["claim_text"]
        evidence_text = self.evidence[evidence_id]

        return (claim_text, evidence_text, label)

```

```

[ ]: class SiameseTripletDataset(Dataset):

    def __init__(
        self,
        claims_paths:List[Path],
        claims_shortlist_paths:List[Path],
        evidence_path:Path,
        evidence_shortlists:List[Path] = None,
        device = None,
        n_neg_shortlist:int = 10,
        n_neg_general:int = 10,
        verbose:bool=True
    ) -> None:
        super(SiameseTripletDataset, self).__init__()
        self.verbose = verbose
        self.device = device
        self.n_neg_shortlist = n_neg_shortlist
        self.n_neg_general = n_neg_general

        # Load claims data from json, this is a list as we could use
        # multiple json files in the same dataset
        self.claims = dict()
        for json_file in claims_paths:
            with open(json_file, mode="r") as f:
                self.claims.update(json.load(fp=f))

```

```

        # print(f"loaded claims: {json_file}")

    # Load the pre-retrieved shortlist of evidences by claim
    self.claims_shortlist = dict()
    for json_file in claims_shortlist_paths:
        with open(json_file, mode="r") as f:
            self.claims_shortlist.update(json.load(fp=f))
            # print(f"loaded claims_shortlist: {json_file}")

    # Load evidence library
    self.evidence = dict()
    with open(evidence_path, mode="r") as f:
        self.evidence.update(json.load(fp=f))
        # print(f"loaded evidences: {json_file}")

    # Load the evidence shortlists if available
    # Reduce the overall evidence list to the shortlist
    if evidence_shortlists is not None:
        self.evidence_shortlist = set()
        for json_file in evidence_shortlists:
            with open(json_file, mode="r") as f:
                self.evidence_shortlist.update(json.load(fp=f))
                # print(f"loaded evidence shortlist: {json_file}")

    # print(f"n_evidences: {len(self.evidence)}")

    # Generate the data
    self.data = self.__generate_data()
    return

def __generate_data(self):
    print("Generate siamese dataset")

    data = []
    for claim_id, claim in tqdm(
        iterable=self.claims.items(),
        desc="claims",
        disable=not self.verbose
    ):
        pos_evidence_ids = set(claim["evidences"])
        shortlist_neg_evidence_ids = \
            set(self.claims_shortlist.get(claim_id, [])) \
            .difference(pos_evidence_ids)
        general_neg_evidence_ids = self.evidence_shortlist

        # For each positive evidence
        for pos_evidence_id in pos_evidence_ids:

```

```

        # Add i number of negatives from shortlisted evidences
        for i in range(self.n_neg_shortlist):
            if len(shortlist_neg_evidence_ids) > 0:
                neg_evidence_ids = random.sample(
                    population=shortlist_neg_evidence_ids,
                    k=min(
                        self.n_neg_shortlist,
                        len(shortlist_neg_evidence_ids)
                    )
                )

                for neg_evidence_id in neg_evidence_ids:
                    data.append(ClaimEvidenceTriple(
                        claim_id=claim_id,
                        pos_evidence_id=pos_evidence_id,
                        neg_evidence_id=neg_evidence_id
                    ))

        # Add j number of negatives from general evidences
        for j in range(self.n_neg_general):
            neg_evidence_ids = random.sample(
                population=general_neg_evidence_ids,
                k=min(self.n_neg_general, len(general_neg_evidence_ids))
            )

            for neg_evidence_id in neg_evidence_ids:
                data.append(ClaimEvidenceTriple(
                    claim_id=claim_id,
                    pos_evidence_id=pos_evidence_id,
                    neg_evidence_id=neg_evidence_id
                ))

        continue

    print(f"Generated data n={len(data)}")

    return data

def __len__(self):
    return len(self.data)

def __getitem__(self, idx) -> Tuple[str]:
    # Fetch the required data rows
    data = self.data[idx]

    # Get text
    anchor_text = self.claims[data.claim_id]["claim_text"]

```

```

pos_evidence_text = self.evidence[data.pos_evidence_id]
neg_evidence_text = self.evidence[data.neg_evidence_id]

return (anchor_text, pos_evidence_text, neg_evidence_text)

```

```
[ ]: # WE WILL GENERATE THE DATASET PER EPOCH SO TO RANDOMISE THE NEGATIVE SAMPLES
```

```

# train_data = SiameseTripletDataset(
#     claims_paths=[DATA_PATH.with_name("train-claims.json")],
#     claims_shortlist_paths=[NER_PATH.
# ↪with_name("train_claim_evidence_retrieved.json")],
#     evidence_shortlists=[NER_PATH.
# ↪with_name("shortlist_train_claim_evidence_retrieved.json")],
#     evidence_path=DATA_PATH.with_name("evidence.json"),
#     device=TORCH_DEVICE,
#     n_neg_shortlist=2,
#     n_neg_general=1
# )

```

```
[ ]: # for i in range(50):
#     anchor, pos, neg = train_data[i]
#     print(anchor[:50], pos[:50], neg[:50])

```

1.3 Build model

```
[ ]: class SiameseTripletEmbedderBert(Module):

    def __init__(
        self,
        pretrained_name:str,
        device,
        **kwargs
    ) -> None:
        super(SiameseTripletEmbedderBert, self).__init__(**kwargs)
        self.device = device

        # Use a pretrained tokenizer
        self.tokenizer = BertTokenizer.from_pretrained(pretrained_name)

        # Use a pretrained model
        self.bert = BertModel.from_pretrained(pretrained_name)
        self.bert.to(device=device)
        return

    def forward(self, anchor_texts, pos_texts, neg_texts=None) -> Tuple[torch.
↪Tensor]:

```

```

# Run the tokenizer
t_kwargs = {
    "return_tensors": "pt",
    "padding": True,
    "truncation": True,
    "max_length": 100,
    "add_special_tokens": True
}
anchor_x = self.tokenizer(anchor_texts, **t_kwargs)
pos_x = self.tokenizer(pos_texts, **t_kwargs)
if neg_texts:
    neg_x = self.tokenizer(neg_texts, **t_kwargs)

anchor_x = anchor_x["input_ids"].to(device=self.device)
pos_x = pos_x["input_ids"].to(device=self.device)
if neg_texts:
    neg_x = neg_x["input_ids"].to(device=self.device)

# Run Bert
anchor_x = self.bert(anchor_x, return_dict=True).pooler_output
pos_x = self.bert(pos_x, return_dict=True).pooler_output
if neg_texts:
    neg_x = self.bert(neg_x, return_dict=True).pooler_output
# dim=768

if neg_texts:
    return anchor_x, pos_x, neg_x
else:
    return anchor_x, pos_x

```

1.4 Training and evaluation loop

```

[ ]: model = SiameseTripletEmbedderBert(
    pretrained_name="bert-base-cased",
    device=TORCH_DEVICE
)

```

Some weights of the model checkpoint at bert-base-cased were not used when initializing BertModel: ['cls.predictions.transform.LayerNorm.bias', 'cls.predictions.bias', 'cls.seq_relationship.weight', 'cls.predictions.transform.dense.bias', 'cls.seq_relationship.bias', 'cls.predictions.decoder.weight', 'cls.predictions.transform.LayerNorm.weight', 'cls.predictions.transform.dense.weight']

- This IS expected if you are initializing BertModel from the checkpoint of a model trained on another task or with another architecture (e.g. initializing a BertForSequenceClassification model from a BertForPreTraining model).
- This IS NOT expected if you are initializing BertModel from the checkpoint of

a model that you expect to be exactly identical (initializing a BertForSequenceClassification model from a BertForSequenceClassification model).

```
[ ]: loss_fn = TripletMarginWithDistanceLoss(
    # distance_function=torch.cdist, # Euclidean distance, not implemented for MPS
    # Default is Pairwise Distance
    margin=10
)
optimizer = Adam(
    params=model.parameters(),
    lr=0.0000002
) #! Hyperparams
```

```
[ ]: run_time = datetime.now().strftime('%Y_%m_%d_%H_%M')
MODEL_NAME = f"model_03a_bert_base_triplet_{run_time}.pth"
N_EPOCHS = 10
BATCH_SIZE = 64
```

```
[ ]: dev_data = SiameseEvalDataset(
    dev_claims_path=DATA_PATH.with_name("dev-claims.json"),
    evidence_path=DATA_PATH.with_name("evidence.json"),
    device=TORCH_DEVICE
)

dev_dataloader = DataLoader(
    dataset=dev_data,
    shuffle=False,
    batch_size=BATCH_SIZE
)
```

```
claims: 100%|          | 154/154 [00:00<00:00, 152772.66it/s]
claims: 100%|          | 154/154 [00:00<00:00, 152772.66it/s]
```

```
[ ]: import warnings
warnings.filterwarnings('ignore')
```

```
[ ]: # Run evaluation before training to establish baseline
model.eval()

dev_batches = tqdm(dev_dataloader, desc="dev batches")
epoch_pos_cos_sim = []
epoch_neg_cos_sim = []
for batch in dev_batches:
    claim_texts, evidence_texts, labels = batch

    # Forward
```

```

claim_emb, evidence_emb = model(claim_texts, evidence_texts)

# Cosine similarity
cos_sim = torch.cosine_similarity(claim_emb, evidence_emb)
labelled_cos_sim = cos_sim * labels
pos_cos_sim = labelled_cos_sim[torch.where(labelled_cos_sim > 0)]
neg_cos_sim = labelled_cos_sim[torch.where(labelled_cos_sim < 0)]

batch_pos_cos_sim = torch.mean(pos_cos_sim).cpu().item()
batch_neg_cos_sim = torch.mean(neg_cos_sim).cpu().item() * -1

epoch_pos_cos_sim.append(batch_pos_cos_sim)
epoch_neg_cos_sim.append(batch_neg_cos_sim)

dev_batches.postfix = f"pos cos_sim: {batch_pos_cos_sim:.3f}" + \
    f" neg cos_sim: {batch_neg_cos_sim:.3f}"

continue

```

```

dev batches: 100%|          | 32/32 [00:08<00:00,  3.64it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

```

```

[ ]: print(f"Average cos sim (pos, neg): {np.mean(epoch_pos_cos_sim):.3f}, {np.
      ↪mean(epoch_neg_cos_sim):.3f}")

```

```

Average cos sim (pos, neg): 0.863419, 0.838396

```

```

[ ]: metric_accuracy = BinaryAccuracy()
metric_f1 = BinaryF1Score()
metric_recall = BinaryF1Score()

scheduler = LinearLR(
    optimizer=optimizer,
    start_factor=0.1,
    end_factor=1,
    total_iters=BATCH_SIZE/10,
    verbose=True
)
best_epoch_loss = 999
for epoch in range(N_EPOCHS):

    print(f"Epoch: {epoch} of {N_EPOCHS}\n")

    # Run training
    model.train()

    train_data = SiameseTripletDataset(

```

```

        claims_paths=[DATA_PATH.with_name("train-claims.json")],
        claims_shortlist_paths=[NER_PATH.
↪with_name("train_claim_evidence_retrieved.json")],
        evidence_shortlists=[NER_PATH.
↪with_name("shortlist_train_claim_evidence_retrieved.json")],
        evidence_path=DATA_PATH.with_name("evidence.json"),
        device=TORCH_DEVICE,
        n_neg_shortlist=2,
        n_neg_general=1
    )

    train_dataloader = DataLoader(
        dataset=train_data,
        shuffle=True,
        batch_size=BATCH_SIZE
    )

    train_batches = tqdm(train_dataloader, desc="train batches")
    running_losses = []
    for batch in train_batches:
        anchor, positive, negative = batch

        # Reset optimizer
        optimizer.zero_grad()

        # Forward + loss
        anchor_emb, pos_emb, neg_emb = model(anchor, positive, negative)
        loss = loss_fn(anchor=anchor_emb, positive=pos_emb, negative=neg_emb)

        # Backward + optimiser
        loss.backward()
        optimizer.step()

        # Update running loss
        batch_loss = loss.item() * len(batch)
        running_losses.append(batch_loss)

        train_batches.postfix = f"loss: {batch_loss:.3f}"

        continue

    scheduler.step()

    epoch_loss = np.average(running_losses)
    print(f"Average epoch loss: {epoch_loss}")

# Save model

```

```

if epoch_loss <= best_epoch_loss:
    best_epoch_loss = epoch_loss
    torch.save(model, MODEL_PATH.with_name(MODEL_NAME))
    print(f"Saved model to: {MODEL_PATH.with_name(MODEL_NAME)}")

# Evaluate every 5 epochs
# if epoch % 5 != 0:
#     continue

# Run evaluation before training to establish baseline
model.eval()

dev_batches = tqdm(dev_dataloader, desc="dev batches")
epoch_pos_cos_sim = []
epoch_neg_cos_sim = []
for batch in dev_batches:
    claim_texts, evidence_texts, labels = batch

    # Forward
    claim_emb, evidence_emb = model(claim_texts, evidence_texts)

    # Cosine similarity
    cos_sim = torch.cosine_similarity(claim_emb, evidence_emb)
    pos_cos_sim = labelled_cos_sim[torch.where(labelled_cos_sim > 0)]
    neg_cos_sim = labelled_cos_sim[torch.where(labelled_cos_sim < 0)]

    batch_pos_cos_sim = torch.mean(pos_cos_sim).cpu().item()
    batch_neg_cos_sim = torch.mean(neg_cos_sim).cpu().item() * -1

    epoch_pos_cos_sim.append(batch_pos_cos_sim)
    epoch_neg_cos_sim.append(batch_neg_cos_sim)

    dev_batches.postfix = f"pos cos_sim: {batch_pos_cos_sim:.3f}" + \
        f" neg cos_sim: {batch_neg_cos_sim:.3f}"

    continue

    print(f"Average cos sim (pos, neg): {np.mean(epoch_pos_cos_sim):3f}, {np.
    ↪mean(epoch_neg_cos_sim):3f}")

print("Done!")

```

Adjusting learning rate of group 0 to 2.0000e-08.
Epoch: 0 of 10

Generate siamese dataset

claims: 100% | 1228/1228 [00:23<00:00, 51.33it/s]

Generated data n=19860

train batches: 100%| | 311/311 [07:02<00:00, 1.36s/it, loss: 29.615]

Adjusting learning rate of group 0 to 4.8125e-08.

Average epoch loss: 29.75464081074264

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.29it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Epoch: 1 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:24<00:00, 50.48it/s]

Generated data n=19860

train batches: 100%| | 311/311 [07:04<00:00, 1.37s/it, loss: 28.745]

Adjusting learning rate of group 0 to 7.6250e-08.

Average epoch loss: 29.271200097258834

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.29it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Epoch: 2 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:24<00:00, 50.37it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:57<00:00, 1.34s/it, loss: 27.977]

Adjusting learning rate of group 0 to 1.0437e-07.

Average epoch loss: 28.525378575279014

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.42it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Epoch: 3 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:23<00:00, 52.97it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:55<00:00, 1.34s/it, loss: 33.199]

Adjusting learning rate of group 0 to 1.3250e-07.

Average epoch loss: 27.53336276732074

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.35it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Epoch: 4 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 53.64it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:54<00:00, 1.33s/it, loss: 26.626]

Adjusting learning rate of group 0 to 1.6062e-07.

Average epoch loss: 23.982970887825036

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.45it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Epoch: 5 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 53.58it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:53<00:00, 1.33s/it, loss: 17.818]

Adjusting learning rate of group 0 to 1.8875e-07.

Average epoch loss: 17.768269468350425

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.42it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Epoch: 6 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 53.51it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:52<00:00, 1.33s/it, loss: 7.174]

Adjusting learning rate of group 0 to 1.8875e-07.

Average epoch loss: 14.633814239041982

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.41it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Epoch: 7 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 53.51it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:49<00:00, 1.32s/it, loss: 13.909]

Adjusting learning rate of group 0 to 1.8875e-07.

Average epoch loss: 12.866380001571017

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.42it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Epoch: 8 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 53.43it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:52<00:00, 1.33s/it, loss: 13.009]

Adjusting learning rate of group 0 to 1.8875e-07.

Average epoch loss: 11.397021522859285

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.45it/s, pos cos_sim: 0.972
neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Epoch: 9 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:23<00:00, 52.66it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:51<00:00, 1.32s/it, loss: 19.942]

Adjusting learning rate of group 0 to 1.8875e-07.

Average epoch loss: 10.945276602671461

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a_bert_base_triplet_2023_05_04_00_12.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.40it/s, pos cos_sim: 0.972

neg cos_sim: 0.868]

Average cos sim (pos, neg): 0.972252, 0.867923

Done!

```
[ ]: @dataclass
class SentenceSimilarity:
    label:int
    score:float
    claim_text:str
    evidence_text:str

[ ]: # Run evaluation before training to establish baseline
model.eval()

dev_batches = tqdm(dev_dataloader, desc="dev batches")
epoch_pos_cos_sim = []
epoch_neg_cos_sim = []
sentence_rankings = []
for batch in dev_batches:
    claim_texts, evidence_texts, labels = batch

    # Forward
    claim_emb, evidence_emb = model(claim_texts, evidence_texts)

    # Cosine similarity
    cos_sim = torch.cosine_similarity(claim_emb, evidence_emb)
    labelled_cos_sim = cos_sim * labels
    pos_cos_sim = labelled_cos_sim[torch.where(labelled_cos_sim > 0)]
    neg_cos_sim = labelled_cos_sim[torch.where(labelled_cos_sim < 0)]

    batch_pos_cos_sim = torch.mean(pos_cos_sim).cpu().item()
    batch_neg_cos_sim = torch.mean(neg_cos_sim).cpu().item() * -1

    epoch_pos_cos_sim.append(batch_pos_cos_sim)
    epoch_neg_cos_sim.append(batch_neg_cos_sim)

for claim_text, evidence_text, label, score in \
```



```

zip(
    claim_texts, evidence_texts,
    labels.detach().cpu().numpy(), cos_sim.detach().cpu().numpy()
):
    sentence_rankings.append(SentenceSimilarity(
        label=label,
        score=score,
        claim_text=claim_text,
        evidence_text=evidence_text
    ))

dev_batches.postfix = f"pos cos_sim: {batch_pos_cos_sim:.3f}" + \
    f" neg cos_sim: {batch_neg_cos_sim:.3f}"

continue

```

```

dev batches: 100%|      | 32/32 [00:07<00:00, 4.41it/s, pos cos_sim: 0.865
neg cos_sim: 0.821]

```

```

[ ]: print(f"Average cos sim (pos, neg): {np.mean(epoch_pos_cos_sim):3f}, {np.
      ↪mean(epoch_neg_cos_sim):3f}")

```

```

Average cos sim (pos, neg): 0.858997, 0.823693

```

```

[ ]: sentence_rankings[:50]

```

```

[ ]: [SentenceSimilarity(label=1, score=0.96096253, claim_text='[South Australia] has
the most expensive electricity in the world.', evidence_text='[citation needed]
South Australia has the highest retail price for electricity in the country. '),
      SentenceSimilarity(label=1, score=0.9753604, claim_text='[South Australia] has
the most expensive electricity in the world.', evidence_text='"South Australia
has the highest power prices in the world."'),
      SentenceSimilarity(label=-1, score=0.82255214, claim_text='[South Australia]
has the most expensive electricity in the world.', evidence_text='"How the
oceans absorb carbon dioxide is critical for predicting climate change."'),
      SentenceSimilarity(label=-1, score=0.77171516, claim_text='[South Australia]
has the most expensive electricity in the world.', evidence_text='Families are
forced into increasing poverty, some facing a daily struggle to pay their rent
and put food on their table. '),
      SentenceSimilarity(label=-1, score=0.88406044, claim_text='[South Australia]
has the most expensive electricity in the world.', evidence_text='They described
the intermediate scenario as the most likely, and that real-world greenhouse gas
forcing had been closest to this scenario. '),
      SentenceSimilarity(label=-1, score=0.9249957, claim_text='[South Australia] has
the most expensive electricity in the world.', evidence_text='Absorption of
infrared light at the vibrational frequencies of atmospheric carbon dioxide
traps energy near the surface, warming the surface and the lower atmosphere. '),

```

SentenceSimilarity(label=-1, score=0.95474094, claim_text='[South Australia] has the most expensive electricity in the world.', evidence_text='Scientists Reach 100% Consensus on Anthropogenic Global Warming.'),

SentenceSimilarity(label=-1, score=0.95091414, claim_text='[South Australia] has the most expensive electricity in the world.', evidence_text='The Labour Party proposed a renegotiation of the withdrawal agreement (towards a closer post-withdrawal relationship with the EU) and would then put this forward as an option in a referendum alongside the option of remaining in the EU.'),

SentenceSimilarity(label=-1, score=0.87609667, claim_text='[South Australia] has the most expensive electricity in the world.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.'),

SentenceSimilarity(label=-1, score=0.9720059, claim_text='[South Australia] has the most expensive electricity in the world.', evidence_text='The 2007 melt season let to a minimum 39% below the 1979-2000 average, and for the first time in human memory, the fabled Northwest Passage opened completely.'),

SentenceSimilarity(label=-1, score=0.8143506, claim_text='[South Australia] has the most expensive electricity in the world.', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.'),

SentenceSimilarity(label=-1, score=0.93211305, claim_text='[South Australia] has the most expensive electricity in the world.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=1, score=0.91068465, claim_text='when 3 per cent of total annual global emissions of carbon dioxide are from humans and Australia prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions reductio\xadn here will have any effect on global climate.', evidence_text='The 2011 UNEP Green Economy report states that "[a]gricultural operations, excluding land use changes, produce approximately 13 per cent of anthropogenic global GHG emissions.'),

SentenceSimilarity(label=1, score=0.86220485, claim_text='when 3 per cent of total annual global emissions of carbon dioxide are from humans and Australia prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions reductio\xadn here will have any effect on global climate.', evidence_text='With a market share of 30% and (potentially) clean electricity, heat pumps could reduce global CO 2 emissions by 8% annually.'),

SentenceSimilarity(label=1, score=0.6428011, claim_text='when 3 per cent of total annual global emissions of carbon dioxide are from humans and Australia prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions reductio\xadn here will have any effect on global climate.', evidence_text='In the modern era, emissions to the atmosphere from volcanoes are approximately 0.645 billion tonnes of CO 2 per year, whereas humans contribute 29 billion tonnes of CO 2 each year.'),

SentenceSimilarity(label=1, score=0.9959356, claim_text='when 3 per cent of

total annual global emissions of carbon dioxide are from humans and Australia
prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions
reductio\xadn here will have any effect on global climate.',
evidence_text='Cumulative anthropogenic (i.e., human-emitted) emissions of CO₂
from fossil fuel use are a major cause of global warming, and give some
indication of which countries have contributed most to human-induced climate
change.')

SentenceSimilarity(label=1, score=0.8453127, claim_text='when 3 per cent of
total annual global emissions of carbon dioxide are from humans and Australia
prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions
reductio\xadn here will have any effect on global climate.',
evidence_text='Other countries with fast growing emissions are South Korea,
Iran, and Australia (which apart from the oil rich Persian Gulf states, now has
the highest percapita emission rate in the world).'),

SentenceSimilarity(label=-1, score=0.65858984, claim_text='when 3 per cent of
total annual global emissions of carbon dioxide are from humans and Australia
prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions
reductio\xadn here will have any effect on global climate.', evidence_text='"How
the oceans absorb carbon dioxide is critical for predicting climate change".'),

SentenceSimilarity(label=-1, score=0.8829576, claim_text='when 3 per cent of
total annual global emissions of carbon dioxide are from humans and Australia
prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions
reductio\xadn here will have any effect on global climate.',
evidence_text='Families are forced into increasing poverty, some facing a daily
struggle to pay their rent and put food on their table.')

SentenceSimilarity(label=-1, score=0.73768455, claim_text='when 3 per cent of
total annual global emissions of carbon dioxide are from humans and Australia
prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions
reductio\xadn here will have any effect on global climate.', evidence_text='They
described the intermediate scenario as the most likely, and that real-world
greenhouse gas forcing had been closest to this scenario.')

SentenceSimilarity(label=-1, score=0.7973865, claim_text='when 3 per cent of
total annual global emissions of carbon dioxide are from humans and Australia
prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions
reductio\xadn here will have any effect on global climate.',
evidence_text='Absorption of infrared light at the vibrational frequencies of
atmospheric carbon dioxide traps energy near the surface, warming the surface
and the lower atmosphere.')

SentenceSimilarity(label=-1, score=0.8495159, claim_text='when 3 per cent of
total annual global emissions of carbon dioxide are from humans and Australia
prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions
reductio\xadn here will have any effect on global climate.',
evidence_text='Scientists Reach 100% Consensus on Anthropogenic Global
Warming.')

SentenceSimilarity(label=-1, score=0.9778296, claim_text='when 3 per cent of
total annual global emissions of carbon dioxide are from humans and Australia
prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions

reductio\xadn here will have any effect on global climate.', evidence_text='The Labour Party proposed a renegotiation of the withdrawal agreement (towards a closer post-withdrawal relationship with the EU) and would then put this forward as an option in a referendum alongside the option of remaining in the EU.')

SentenceSimilarity(label=-1, score=0.72408855, claim_text='when 3 per cent of total annual global emissions of carbon dioxide are from humans and Australia prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions reductio\xadn here will have any effect on global climate.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.')

SentenceSimilarity(label=-1, score=0.96708447, claim_text='when 3 per cent of total annual global emissions of carbon dioxide are from humans and Australia prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions reductio\xadn here will have any effect on global climate.', evidence_text='The 2007 melt season let to a minimum 39% below the 1979-2000 average, and for the first time in human memory, the fabled Northwest Passage opened completely.')

SentenceSimilarity(label=-1, score=0.6464118, claim_text='when 3 per cent of total annual global emissions of carbon dioxide are from humans and Australia prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions reductio\xadn here will have any effect on global climate.', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.')

SentenceSimilarity(label=-1, score=0.8081426, claim_text='when 3 per cent of total annual global emissions of carbon dioxide are from humans and Australia prod\xaduces 1.3 per cent of this 3 per cent, then no amount of emissions reductio\xadn here will have any effect on global climate.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.')

SentenceSimilarity(label=1, score=0.995721, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='Multiple independently produced instrumental datasets confirm that the 2009-2018 decade was 0.93 ± 0.07 \xa0°C warmer than the pre-industrial baseline (1850-1900).')

SentenceSimilarity(label=1, score=0.99781245, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='The planet is now 0.8\xa0°C warmer than in pre-industrial times.')

SentenceSimilarity(label=-1, score=0.98314357, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='"How the oceans absorb carbon dioxide is critical for predicting climate change.')

SentenceSimilarity(label=-1, score=0.4634062, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='Families are forced into increasing poverty, some facing a daily struggle to pay their rent and put food on their table.')

SentenceSimilarity(label=-1, score=0.99332356, claim_text='This means that the

world is now 1C warmer than it was in pre-industrial times', evidence_text='They described the intermediate scenario as the most likely, and that real-world greenhouse gas forcing had been closest to this scenario.')

SentenceSimilarity(label=-1, score=0.9807402, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='Absorption of infrared light at the vibrational frequencies of atmospheric carbon dioxide traps energy near the surface, warming the surface and the lower atmosphere.')

SentenceSimilarity(label=-1, score=0.94999415, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='Scientists Reach 100% Consensus on Anthropogenic Global Warming.')

SentenceSimilarity(label=-1, score=0.7601046, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='The Labour Party proposed a renegotiation of the withdrawal agreement (towards a closer post-withdrawal relationship with the EU) and would then put this forward as an option in a referendum alongside the option of remaining in the EU.')

SentenceSimilarity(label=-1, score=0.9954001, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.')

SentenceSimilarity(label=-1, score=0.8038523, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='The 2007 melt season let to a minimum 39% below the 1979-2000 average, and for the first time in human memory, the fabled Northwest Passage opened completely.')

SentenceSimilarity(label=-1, score=0.9855043, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.')

SentenceSimilarity(label=-1, score=0.9774, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.')

SentenceSimilarity(label=1, score=0.7316737, claim_text='"As it happens, Zika may also be a good model of the second worrying effect - disease mutation.', evidence_text='Genetic disorders are the result of deleterious mutations and can be due to spontaneous mutation in the affected individual, or can be inherited.')

SentenceSimilarity(label=1, score=0.92852515, claim_text='"As it happens, Zika may also be a good model of the second worrying effect - disease mutation.', evidence_text='These errors, called mutations, can affect the phenotype of an organism, especially if they occur within the protein coding sequence of a gene.')

SentenceSimilarity(label=1, score=0.98385155, claim_text='"As it happens, Zika

may also be a good model of the second worrying effect - disease mutation.',
evidence_text='This is especially useful studying diseases in adults by allowing
expression after a certain period of growth, thus eliminating the deleterious
effect of gene expression seen during stages of development in model
organisms.')

SentenceSimilarity(label=1, score=0.9889933, claim_text='"As it happens, Zika
may also be a good model of the second worrying effect - disease mutation.',
evidence_text='In addition, research was needed to determine the mechanism by
which Zika produced these effects.')

SentenceSimilarity(label=1, score=0.9377308, claim_text='"As it happens, Zika
may also be a good model of the second worrying effect - disease mutation.',
evidence_text='Zika also appears to have an equal tropism for cells of the
developing eye, leading to high rates of eye abnormalities as well.')

SentenceSimilarity(label=-1, score=0.76192385, claim_text='"As it happens, Zika
may also be a good model of the second worrying effect - disease mutation.',
evidence_text='"How the oceans absorb carbon dioxide is critical for predicting
climate change".'),

SentenceSimilarity(label=-1, score=0.83162403, claim_text='"As it happens, Zika
may also be a good model of the second worrying effect - disease mutation.',
evidence_text='Families are forced into increasing poverty, some facing a daily
struggle to pay their rent and put food on their table.')

SentenceSimilarity(label=-1, score=0.83066607, claim_text='"As it happens, Zika
may also be a good model of the second worrying effect - disease mutation.',
evidence_text='They described the intermediate scenario as the most likely, and
that real-world greenhouse gas forcing had been closest to this scenario.')

SentenceSimilarity(label=-1, score=0.8755104, claim_text='"As it happens, Zika
may also be a good model of the second worrying effect - disease mutation.',
evidence_text='Absorption of infrared light at the vibrational frequencies of
atmospheric carbon dioxide traps energy near the surface, warming the surface
and the lower atmosphere.')

SentenceSimilarity(label=-1, score=0.903653, claim_text='"As it happens, Zika
may also be a good model of the second worrying effect - disease mutation.',
evidence_text='Scientists Reach 100% Consensus on Anthropogenic Global
Warming.')

SentenceSimilarity(label=-1, score=0.9699, claim_text='"As it happens, Zika may
also be a good model of the second worrying effect - disease mutation.',
evidence_text='The Labour Party proposed a renegotiation of the withdrawal
agreement (towards a closer post-withdrawal relationship with the EU) and would
then put this forward as an option in a referendum alongside the option of
remaining in the EU.'])]

```
[ ]: pos_sentence_rankings = sorted([
    s for s in sentence_rankings
    if s.label == 1
], key=lambda s: s.score)

neg_sentence_rankings = sorted([
```

```

    s for s in sentence_rankings
    if s.label == -1
], key=lambda s: s.score, reverse=True)

```

```
[ ]: pos_sentence_rankings[:50]
```

```
[ ]: [SentenceSimilarity(label=1, score=0.1249495, claim_text='In many other cases,
though - hurricanes, for example - the linkage to global warming for particular
trends is uncertain or disputed.', evidence_text='Although there are a few areas
of linkage, the relationship between the two is weak.'),
SentenceSimilarity(label=1, score=0.39303237, claim_text='More importantly, the
OISM list only contains 39 scientists who specialise in climate science.',
evidence_text='A botanist, plant scientist or phytologist is a scientist who
specialises in this field.'),
SentenceSimilarity(label=1, score=0.43102425, claim_text="'Another global
warming myth comes crashing down.", evidence_text='Delingpole has engaged in
climate change denialism; in 2009 he wrote of "The conspiracy behind the
Anthropogenic Global Warming myth.'),
SentenceSimilarity(label=1, score=0.43921182, claim_text='More importantly, the
OISM list only contains 39 scientists who specialise in climate science.',
evidence_text='The list includes scientists from several specialities or
disciplines.'),
SentenceSimilarity(label=1, score=0.47019237, claim_text="'Which is to say that
these beans will be eaten by cows, and the cows will convert the beans to meat,
and the humans will eat the meat.', evidence_text='are vehemently against the
consumption of meat and eggs (though they do consume and encourage the
consumption of milk, butter and cheese.'),
SentenceSimilarity(label=1, score=0.50972927, claim_text='The gas builds up in
the soil, forming mounds called 'pingoes.'', evidence_text='At the next larger
scale, soil structures called peds or more commonly soil aggregates are created
from the soil separates when iron oxides, carbonates, clay, silica and humus,
coat particles and cause them to adhere into larger, relatively stable secondary
structures.'),
SentenceSimilarity(label=1, score=0.51082236, claim_text='Satellite
measurements of infrared spectra over the past 40 years observe less energy
escaping to space at the wavelengths associated with CO2.', evidence_text='The
three channels use the same frequency but different carbon dioxide cell
pressure, the corresponding weighting functions peaks at 29\xa0km for channel 1,
37\xa0km for channel 2 and 45\xa0km for channel 3.'),
SentenceSimilarity(label=1, score=0.5120701, claim_text='Global Warming history
completely coincides with the history of artificial satellites and the use of
microwave frequencies from outer space.', evidence_text='A communications
satellite is an artificial satellite that relays and amplifies radio
telecommunications signals via a transponder; it creates a communication channel
between a source transmitter and a receiver at different locations on Earth.'),
SentenceSimilarity(label=1, score=0.51591307, claim_text="'Another global
warming myth comes crashing down.", evidence_text='The conspiracy behind the
```

Anthropogenic Global Warming myth (aka AGW; aka ManBearPig) has been suddenly, brutally and quite deliciously exposed", and "A contretemps with a Climate Bully who wonders whether I have a science degree.'),

SentenceSimilarity(label=1, score=0.5348499, claim_text='Jim Hansen had several possible scenarios; his mid-level scenario B was right.', evidence_text='An alternate scenario would keep the warming to below this if climate sensitivity were below 3\xa0°C for doubled CO 2.'),

SentenceSimilarity(label=1, score=0.545984, claim_text='"Which is to say that these beans will be eaten by cows, and the cows will convert the beans to meat, and the humans will eat the meat.', evidence_text='Buffalo are generally fed on coarse feeds; they convert them into the protein-rich lean meat.'),

SentenceSimilarity(label=1, score=0.54724556, claim_text='Newt Gingrich "teamed with Nancy Pelosi and Al Gore on global warming."', evidence_text='He was known as one of the Atari Democrats, later called the "Democrats\' Greens, politicians who see issues like clean air, clean water and global warming as the key to future victories for their party."'),

SentenceSimilarity(label=1, score=0.5524541, claim_text='But [climate scientists] say that aspects of the case of Hurricane Harvey-and the recent history of tropical cyclones worldwide-suggest global warming is making a bad situation worse.', evidence_text='Global warming not only causes changes in tropical cyclones, it may also make some impacts from them worse via sea level rise.'),

SentenceSimilarity(label=1, score=0.5646414, claim_text='With 32 years of rapidly increasing global temperatures and only a minor increase in global CO2 emissions, followed by 33 years of slowly cooling global temperatures with rapid increases in global CO2 emissions, it was deceitful for the IPCC to make any claim that CO2 emissions were primarily responsible for observed 20th century global warming."', evidence_text='The global warming observed over the past 50 years is due primarily to human-induced emissions of heat-trapping gases.'),

SentenceSimilarity(label=1, score=0.57007754, claim_text='many scientists were surprised when other researchers subsequently found that ringed and bearded seals (the primary prey of polar bears) north of the Bering Strait especially thrived with a longer open-water season, which is particularly conducive to fishing', evidence_text='The polar bear is the most carnivorous member of the bear family, and throughout most of its range, its diet primarily consists of ringed (Pusa hispida) and bearded seals (Erignathus barbatus).'),

SentenceSimilarity(label=1, score=0.57023895, claim_text='Many of the world's coral reefs are already barren or in a state of constant decline.', evidence_text='For example, Midway Atoll in Hawaii supports nearly three million seabirds, including two-thirds (1.5 million) of the global population of Laysan albatross, and one-third of the global population of black-footed albatross.'),

SentenceSimilarity(label=1, score=0.5789369, claim_text='[Riebesell] is a world authority on the topic and has typically communicated cautiously about the effects of acidification.', evidence_text='Perhaps one of the most recent adverse effects of climate change to be explored is that of ocean acidification.'),

SentenceSimilarity(label=1, score=0.57993543, claim_text='The unlikely scenarios are now, all of a sudden, becoming more probable than they once were thought to be,' says Sweet.', evidence_text='The reverse possibility, that a small group broke off and wandered from India into Western Asia is readily dismissed as an improbably long migration, again without the least bit of evidence.'),

SentenceSimilarity(label=1, score=0.5865245, claim_text='"Austria is today seeing its earliest snowfall in history with 30 to 40 centimetres already predicted in the mountains.', evidence_text="This law of neutrality, passed in late October 1955 (and not the State Treaty itself), ensured that modern Austria would align with neither NATO nor the Soviet bloc, and is considered one of the reasons for Austria's delayed entry into the European Union in 1995."),

SentenceSimilarity(label=1, score=0.58949006, claim_text="Al Gore's book is quite accurate, and far more accurate than contrarian books.", evidence_text='"He is making a qualitative point, which is entirely accurate.'),

SentenceSimilarity(label=1, score=0.59098095, claim_text='The heatwave we now have in Europe is not something that was expected with just 1C of warming', evidence_text="2018's temperature was 16.1\xa0°C (61.0\xa0°F), meaning it ranks as the 18th warmest June recorded in England in the past 359 years, also being the warmest since 1976."),

SentenceSimilarity(label=1, score=0.596289, claim_text='In recent decades this warming has been accompanied by a constant rise in the sea level and, it would appear, by an increase of extreme weather events, even if a scientifically determinable cause cannot be assigned to each particular phenomenon.', evidence_text='Between 1993 and 2017, the global mean sea level rose on average by 3.1 ±\xa00.3\xa0mm per year, with an acceleration detected as well.'),

SentenceSimilarity(label=1, score=0.59978306, claim_text='Jim Hansen had several possible scenarios; his mid-level scenario B was right.', evidence_text='They described a business-as-usual scenario, which has greenhouse gases growing at approximately 2% per year; and an alternate scenario, in which greenhouse gases concentrations decline.'),

SentenceSimilarity(label=1, score=0.6011211, claim_text='many scientists were surprised when other researchers subsequently found that ringed and bearded seals (the primary prey of polar bears) north of the Bering Strait especially thrived with a longer open-water season, which is particularly conducive to fishing', evidence_text='The polar bear is the most carnivorous species of bear, and its diet primarily consists of ringed (Pusa hispida) and bearded (Erignathus barbatus) seals.'),

SentenceSimilarity(label=1, score=0.6011776, claim_text='\To suddenly label CO2 as a "pollutant" is a disservice to a gas that has played an enormous role in the development and sustainability of all life on this wonderful Earth.', evidence_text='Emissions trading (also known as cap and trade) is a market-based approach to controlling pollution by providing economic incentives for achieving reductions in the emissions of pollutants.'),

SentenceSimilarity(label=1, score=0.6042464, claim_text='Continued greenhouse gas emissions at or above current rates would cause further warming and induce many changes in the global climate system during the 21st century that would

very likely be larger than those observed during the 20th century.',
evidence_text='Many regions have probably already seen increases in warm spells and heat waves, and it is virtually certain that these changes will continue over the 21st century.'),

SentenceSimilarity(label=1, score=0.6089132, claim_text='But each serial adjustment has tended to make the early years colder, which increases the warming trend.', evidence_text='Climate change adaptation is "the adjustment in natural or human systems in response to actual or expected climatic stimuli or their effects, which moderates harm or exploits beneficial opportunities".'),

SentenceSimilarity(label=1, score=0.6143447, claim_text='In recent decades this warming has been accompanied by a constant rise in the sea level and, it would appear, by an increase of extreme weather events, even if a scientifically determinable cause cannot be assigned to each particular phenomenon.', evidence_text='The effects of global warming include rising sea levels, regional changes in precipitation, more frequent extreme weather events such as heat waves, and expansion of deserts.'),

SentenceSimilarity(label=1, score=0.61719227, claim_text='"Moreover, the ocean already contains so-called oxygen minimum zones, generally found in the middle depths.', evidence_text='These areas are also known as oxygen minimum zones (OMZ).'),

SentenceSimilarity(label=1, score=0.61810946, claim_text='The particular signature of warming in 2016 was also revealing in another way, Overpeck said, noting that the stratosphere... saw record cold temperatures last year', evidence_text='One of the strongest predictions of the greenhouse effect is that the stratosphere will cool.'),

SentenceSimilarity(label=1, score=0.62355196, claim_text='While there has been a mean rise of a little more than 3mm per year worldwide since the 1990s, in the last decade, the NOAA Virginia Key tide gauge just south of Miami Beach has measured a 9mm rise annually."', evidence_text='Between 1993 and 2017, the global mean sea level rose on average by 3.1 ±\xa00.3\xa0mm per year, with an acceleration detected as well.'),

SentenceSimilarity(label=1, score=0.62389886, claim_text='The claim that 97 percent of scientists believe humans are causing climate change has been debunked by the "head" of the United Nations Intergovernmental Panel on Climate Change.', evidence_text='The study found that 97% of the 489 scientists surveyed agreed that global temperatures have risen over the past century.'),

SentenceSimilarity(label=1, score=0.6240232, claim_text="Without the forests' humidity, previously moisture-laden winds blew dry.", evidence_text='Laurel forest and other cloud forest are subtropical and mild temperate forest, found in areas with high humidity and relatively stable and mild temperatures.'),

SentenceSimilarity(label=1, score=0.6309831, claim_text='"Twentieth century global warming did not start until 1910.', evidence_text='The average global temperature on Earth has increased by about 0.8° Celsius (1.4° Fahrenheit) since 1880; Two-thirds of the warming has occurred since 1975, at a rate of roughly 0.15-0.20\xa0°C per decade.'),

SentenceSimilarity(label=1, score=0.64097816, claim_text='The grasslands, crops, forests and territorial waters of Australia absorb more carbon dioxide

than Australia emits.', evidence_text='It is estimated that forests absorb between 10 and 20 tons of carbon dioxide per hectare each year, through photosynthetic conversion into starch, cellulose, lignin, and other components of wooden biomass.'),

SentenceSimilarity(label=1, score=0.6413591, claim_text='\'To suddenly label CO2 as a "pollutant" is a disservice to a gas that has played an enormous role in the development and sustainability of all life on this wonderful Earth.', evidence_text='In its book, the Commission described four key elements of sustainability with respect to energy: the ability to increase the supply of energy to meet growing human needs, energy efficiency and conservation, public health and safety, and "protection of the biosphere and prevention of more localized forms of pollution."'),

SentenceSimilarity(label=1, score=0.6428011, claim_text='when 3 per cent of total annual global emissions of carbon dioxide are from humans and Australia produces 1.3 per cent of this 3 per cent, then no amount of emissions reduction here will have any effect on global climate.', evidence_text='In the modern era, emissions to the atmosphere from volcanoes are approximately 0.645 billion tonnes of CO2 per year, whereas humans contribute 29 billion tonnes of CO2 each year.'),

SentenceSimilarity(label=1, score=0.6454537, claim_text='\"Which is to say that these beans will be eaten by cows, and the cows will convert the beans to meat, and the humans will eat the meat.', evidence_text='conversion of human-inedible residues of food crops.'),

SentenceSimilarity(label=1, score=0.65332663, claim_text='Thick arctic sea ice is in rapid retreat.', evidence_text='In 2002 the 12\xa0km (7.5\xa0mi) long floating terminus of the glacier entered a phase of rapid retreat, with the ice front breaking up and the floating terminus disintegrating and accelerating to a retreat rate of over 30\xa0m (98\xa0ft) per day.'),

SentenceSimilarity(label=1, score=0.6544106, claim_text='\"Multiple lines of evidence indicate Greenland's ice loss is accelerating and will contribute sea level rise in the order of metres over the next few centuries.', evidence_text='The glaciers of Greenland are also contributing to a rise in the global sea level faster than was previously believed.'),

SentenceSimilarity(label=1, score=0.663985, claim_text='The claim that 97 percent of scientists believe humans are causing climate change has been debunked by the "head" of the United Nations Intergovernmental Panel on Climate Change.', evidence_text='97% of the scientists surveyed agreed that global temperatures had increased during the past 100 years; 84% said they personally believed human-induced warming was occurring, and 74% agreed that "currently available scientific evidence" substantiated its occurrence.'),

SentenceSimilarity(label=1, score=0.6649524, claim_text='Protecting and restoring forests would reduce 18% of emissions by 2030', evidence_text='Japan needs to reduce 26% of green house gas emission from 2013 by 2030 to accomplish Paris Agreement and is trying to reduce 2% of them by forestry.'),

SentenceSimilarity(label=1, score=0.6679884, claim_text='Jim Hansen had several possible scenarios; his mid-level scenario B was right.', evidence_text='They described the intermediate scenario as the most likely, and that real-world

```

greenhouse gas forcing had been closest to this scenario. '),
SentenceSimilarity(label=1, score=0.67599887, claim_text='many scientists were
surprised when other researchers subsequently found that ringed and bearded
seals (the primary prey of polar bears) north of the Bering Strait especially
thrived with a longer open-water season, which is particularly conducive to
fishing', evidence_text='Ringed seals are one of the primary prey of polar bears
and killer whales, and have long been a component of the diet of indigenous
people of the Arctic. '),
SentenceSimilarity(label=1, score=0.67722476, claim_text='The unlikely
scenarios are now, all of a sudden, becoming more probable than they once were
thought to be,' says Sweet.'', evidence_text="In recent decades, a few
specialists have continued to support this interpretation, and Peter Schrijver
has said that 'to a large extent, it is linguistics that is responsible for
thinking in terms of drastic scenarios' about demographic change in late Roman
Britain." ),
SentenceSimilarity(label=1, score=0.67887986, claim_text='many scientists were
surprised when other researchers subsequently found that ringed and bearded
seals (the primary prey of polar bears) north of the Bering Strait especially
thrived with a longer open-water season, which is particularly conducive to
fishing', evidence_text='Natural predators of the bearded seal include polar
bears, who rely on these seals as a major food source. '),
SentenceSimilarity(label=1, score=0.6792201, claim_text='"Austria is today
seeing its earliest snowfall in history with 30 to 40 centimetres already
predicted in the mountains.', evidence_text="Austria lies between latitudes 46°
and 49° N, and longitudes 9° and 18° E. It can be divided into five areas, the
biggest being the Eastern Alps, which constitute 62% of the nation's total
area." ),
SentenceSimilarity(label=1, score=0.6804468, claim_text='In the past, warming
has never been a threat to life on Earth.', evidence_text='Given the potential
threat to marine ecosystems and its ensuing impact on human society and economy,
especially as it acts in conjunction with anthropogenic global warming, there is
an urgent need for immediate action.''),
SentenceSimilarity(label=1, score=0.6827507, claim_text='Many of the world's
coral reefs are already barren or in a state of constant decline.',
evidence_text='Aquaculture is showing promise as a potentially effective tool
for restoring coral reefs, which have been declining around the world.'),
SentenceSimilarity(label=1, score=0.68677354, claim_text='Water vapor helps
trap heat, and is a far the strongest of the major greenhouse gases,
contributing 36-72 percent of the greenhouse effect.', evidence_text='From 1859
onwards, he showed that the effect was due to a very small proportion of the
atmosphere, with the main gases having no effect, and was largely due to water
vapour, though small percentages of hydrocarbons and carbon dioxide had a
significant effect.'')]

```

```
[ ]: neg_sentence_rankings[:50]
```

[]: [SentenceSimilarity(label=-1, score=0.9969016, claim_text='The Earth's climate is changing in response to increasing concentrations of greenhouse gases (GHGs) and particulate matter in the atmosphere, largely as the result of human activities.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.99676955, claim_text='Once natural influences, in particular the impact of El Niño and La Niña, are removed from the recent temperature record, there is no evidence of a significant change in the human contribution to climate change.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.'),

SentenceSimilarity(label=-1, score=0.9967185, claim_text="Multiple lines of evidence indicate Greenland's ice loss is accelerating and will contribute sea level rise in the order of metres over the next few centuries.", evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.'),

SentenceSimilarity(label=-1, score=0.9966046, claim_text="There are a myriad of other radiative forcings that affect the planet's energy imbalance.", evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.'),

SentenceSimilarity(label=-1, score=0.996209, claim_text='[Ice] is expanding in much of Antarctica, contrary to the widespread public belief that global warming is melting the continental ice cap.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.'),

SentenceSimilarity(label=-1, score=0.9959967, claim_text='Evidence is building that net cloud feedback is likely positive and unlikely to be strongly negative.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.99590933, claim_text='In many other cases, though - hurricanes, for example - the linkage to global warming for particular trends is uncertain or disputed.', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.'),

SentenceSimilarity(label=-1, score=0.9954001, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.'),

SentenceSimilarity(label=-1, score=0.99539846, claim_text='By 2050 there's a

scientific consensus that we reached the tipping point for ice sheets in Greenland and the West Antarctic', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.')

SentenceSimilarity(label=-1, score=0.9953644, claim_text='no one really knows if last year 2016 was a global temperature record.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.')

SentenceSimilarity(label=-1, score=0.99528664, claim_text='IPCC overestimate temperature rise.', evidence_text='"How the oceans absorb carbon dioxide is critical for predicting climate change".'),

SentenceSimilarity(label=-1, score=0.9952617, claim_text='While the greenhouse effect is a natural occurrence, too much warming has severe negative impacts on agriculture, health and environment.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.')

SentenceSimilarity(label=-1, score=0.99522245, claim_text='Some global warming "skeptics" argue that the Earth's climate sensitivity is so low that a doubling of atmospheric CO2 will result in a surface temperature change on the order of 1°C or less, and that therefore global warming is nothing to worry about.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.')

SentenceSimilarity(label=-1, score=0.9950216, claim_text='Lake-bottom sediments in Florida tell us that recent major hurricane activity in the Gulf of Mexico has been less frequent than in centuries past.', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.')

SentenceSimilarity(label=-1, score=0.994888, claim_text='Volcanoes emit around 0.3 billion tonnes of CO2 per year.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.')

SentenceSimilarity(label=-1, score=0.99479485, claim_text='The team of climate scientists notes that in failing to predict the warming "hiatus" in the beginning of the 21st century, the Intergovernmental Panel on Climate Change (IPCC) models overestimated temperature increases...', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.')

SentenceSimilarity(label=-1, score=0.99448144, claim_text='While there are isolated cases of growing glaciers, the overwhelming trend in glaciers worldwide is retreat.', evidence_text='Absorption of infrared light at the vibrational frequencies of atmospheric carbon dioxide traps energy near the surface, warming the surface and the lower atmosphere.')

SentenceSimilarity(label=-1, score=0.9944575, claim_text='Many of the world's

coral reefs are already barren or in a state of constant decline.',
evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years. '),
SentenceSimilarity(label=-1, score=0.99429655, claim_text='Clouds provide negative feedback', evidence_text='Scientists Reach 100% Consensus on Anthropogenic Global Warming. '),
SentenceSimilarity(label=-1, score=0.9941132, claim_text='(Kerr 2007) points out that the sunlight-reflecting haze that cools much of the planet seems to have thinned over the past decade or so.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation. '),
SentenceSimilarity(label=-1, score=0.9937247, claim_text='CO2 limits won't cool the planet.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years. '),
SentenceSimilarity(label=-1, score=0.9936684, claim_text='The long term trend from albedo is that of cooling.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation. '),
SentenceSimilarity(label=-1, score=0.9934684, claim_text='Once natural influences, in particular the impact of El Niño and La Niña, are removed from the recent temperature record, there is no evidence of a significant change in the human contribution to climate change.', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation. '),
SentenceSimilarity(label=-1, score=0.9933597, claim_text='While the greenhouse effect is a natural occurrence, too much warming has severe negative impacts on agriculture, health and environment.', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation. '),
SentenceSimilarity(label=-1, score=0.99332786, claim_text='The IPCC simply updated their temperature history graphs to show the best data available at the time.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years. '),
SentenceSimilarity(label=-1, score=0.99332356, claim_text='This means that the world is now 1C warmer than it was in pre-industrial times', evidence_text='They described the intermediate scenario as the most likely, and that real-world greenhouse gas forcing had been closest to this scenario. '),
SentenceSimilarity(label=-1, score=0.9931571, claim_text='The noted oceanographer Walter Munk referred to sea-level rise as an "enigma"', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years. '),

SentenceSimilarity(label=-1, score=0.9931505, claim_text='Around 97% of climate experts agree that humans are causing global warming.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.9931032, claim_text='While there are isolated cases of growing glaciers, the overwhelming trend in glaciers worldwide is retreat.', evidence_text='Scientists Reach 100% Consensus on Anthropogenic Global Warming.'),

SentenceSimilarity(label=-1, score=0.9930634, claim_text='The Earth's climate is changing in response to increasing concentrations of greenhouse gases (GHGs) and particulate matter in the atmosphere, largely as the result of human activities.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.'),

SentenceSimilarity(label=-1, score=0.9930149, claim_text='Greg Hunt CSIRO research shows carbon emissions can be reduced by 20 per cent over 40 years using nature, soils and trees.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.9929757, claim_text='Greenland has only lost a tiny fraction of its ice mass', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.'),

SentenceSimilarity(label=-1, score=0.99290794, claim_text='The science is clear, climate change is making extreme weather events, including tornadoes, worse.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.'),

SentenceSimilarity(label=-1, score=0.9928749, claim_text='"Global warming' and 'climate change' mean different things and have both been used for decades.", evidence_text='They described the intermediate scenario as the most likely, and that real-world greenhouse gas forcing had been closest to this scenario.'),

SentenceSimilarity(label=-1, score=0.9928216, claim_text='While there are isolated cases of growing glaciers, the overwhelming trend in glaciers worldwide is retreat.', evidence_text='Projections based on the Special Report on Emissions Scenarios suggest warming over the 21st century at a more rapid rate than that experienced for at least the last 10,000 years.'),

SentenceSimilarity(label=-1, score=0.9927574, claim_text='This means the jet stream meanders more, with big loops bringing warm air to the frozen north and cold air into warmer southern climes."', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.9925295, claim_text='We have had ice ages and warmer periods when alligators were found in Spitzbergen.',

evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.992389, claim_text='Around 97% of climate experts agree that humans are causing global warming.', evidence_text='They described the intermediate scenario as the most likely, and that real-world greenhouse gas forcing had been closest to this scenario.'),

SentenceSimilarity(label=-1, score=0.9923562, claim_text='The CO2 amplifies the warming and mixes through the atmosphere, spreading warming throughout the planet.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.99228823, claim_text='Clouds provide negative feedback', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.9922634, claim_text='IPCC report warning last week the world is "nowhere near on track" to meet its Paris commitments', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.9921982, claim_text='The actual data show high northern latitudes are warmer today than in 1940.', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.'),

SentenceSimilarity(label=-1, score=0.9919982, claim_text='Thick arctic sea ice is in rapid retreat.', evidence_text='Surface heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation.'),

SentenceSimilarity(label=-1, score=0.9917964, claim_text='The IPCC simply updated their temperature history graphs to show the best data available at the time.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.9917532, claim_text='Heat is continuing to build up in the subsurface ocean.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation.'),

SentenceSimilarity(label=-1, score=0.99157894, claim_text='Water vapor helps trap heat, and is a far the strongest of the major greenhouse gases, contributing 36-72 percent of the greenhouse effect.', evidence_text='Surface

heat and freshwater fluxes create global density gradients that drive the thermohaline circulation part of large-scale ocean circulation. '),

SentenceSimilarity(label=-1, score=0.99154854, claim_text='Thick arctic sea ice is in rapid retreat.', evidence_text='How the oceans absorb carbon dioxide is critical for predicting climate change'. '),

SentenceSimilarity(label=-1, score=0.9915379, claim_text='The science is clear, climate change is making extreme weather events, including tornadoes, worse.', evidence_text='They described the intermediate scenario as the most likely, and that real-world greenhouse gas forcing had been closest to this scenario. '),

SentenceSimilarity(label=-1, score=0.9915007, claim_text='Austria is today seeing its earliest snowfall in history with 30 to 40 centimetres already predicted in the mountains.', evidence_text='The change was attributed to increasingly cold winters in the Arctic stratosphere at an altitude of approximately 20\xa0km (12\xa0mi), a change associated with global warming in a relationship that is still under investigation. '),

SentenceSimilarity(label=-1, score=0.9914047, claim_text='Clouds provide negative feedback', evidence_text='Absorption of infrared light at the vibrational frequencies of atmospheric carbon dioxide traps energy near the surface, warming the surface and the lower atmosphere. ')]