model_03a_bert_base_triplet_margin_1

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

```python
# 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

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

1.1.3 Dependencies

```python
# 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

```python
@dataclass
class ClaimEvidencePair:
    claim_id:str
    evidence_id:str
    label:int = 0
```

```python
@dataclass
class ClaimEvidenceTriple:
    claim_id:str
    pos_evidence_id:str
    neg_evidence_id:str
```

```python
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))
```

2

```python
        # 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
```

```python
        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)
```

```python
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))
```

```python
            # 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:
```

```python
            # 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]:
```

```python
        # 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

```python
[ ]: 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.weight',
'cls.predictions.transform.LayerNorm.bias',
'cls.predictions.transform.dense.weight',
'cls.predictions.transform.dense.bias', 'cls.seq_relationship.bias',
'cls.predictions.decoder.weight', 'cls.predictions.bias',
'cls.seq_relationship.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).

```python
loss_fn = TripletMarginWithDistanceLoss(
    # distance_function=torch.cdist, # Euclidean distance, not implemented for␣
  ↪MPS
    # Default is Pairwise Distance
    margin=1
)
optimizer = Adam(
    params=model.parameters(),
    lr=0.0000002
) #! Hyperparams
```

```python
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
```

```python
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, 181337.12it/s]
claims: 100%|       | 154/154 [00:00<00:00, 181337.12it/s]
```

```python
import warnings
warnings.filterwarnings('ignore')
```

```python
# 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
```

```python
    # 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]
```

```python
[ ]: 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
```

```python
[ ]: 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()
```

```python
    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}")
```

```python
        # 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)
            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

    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 [00:24<00:00, 50.42it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:59<00:00, 1.35s/it, loss: 8.485]

Adjusting learning rate of group 0 to 4.8125e-08.
Average epoch loss: 6.6321448495916995
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.32it/s, pos cos_sim: 0.977
neg cos_sim: 0.904]

Average cos sim (pos, neg): 0.899169, 0.884421
Epoch: 1 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:23<00:00, 51.79it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:51<00:00, 1.32s/it, loss: 2.821]

Adjusting learning rate of group 0 to 7.6250e-08.
Average epoch loss: 5.160400438730356
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.38it/s, pos cos_sim: 0.981
neg cos_sim: 0.957]

Average cos sim (pos, neg): 0.950285, 0.947763
Epoch: 2 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 53.94it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:54<00:00, 1.33s/it, loss: 3.469]

Adjusting learning rate of group 0 to 1.0437e-07.
Average epoch loss: 3.6845154101058983
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.31it/s, pos cos_sim: 0.984
neg cos_sim: 0.982]

Average cos sim (pos, neg): 0.976250, 0.976166
Epoch: 3 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 53.42it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:54<00:00, 1.33s/it, loss: 2.515]

Adjusting learning rate of group 0 to 1.3250e-07.
Average epoch loss: 2.976541721935824
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.36it/s, pos cos_sim: 0.990
neg cos_sim: 0.986]

Average cos sim (pos, neg): 0.982831, 0.981949
Epoch: 4 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 54.06it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:51<00:00, 1.32s/it, loss: 2.231]

Adjusting learning rate of group 0 to 1.6062e-07.
Average epoch loss: 2.4536364891138107
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.39it/s, pos cos_sim: 0.989
neg cos_sim: 0.985]

Average cos sim (pos, neg): 0.983680, 0.981273
Epoch: 5 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 54.06it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:50<00:00, 1.32s/it, loss: 2.354]

Adjusting learning rate of group 0 to 1.8875e-07.
Average epoch loss: 1.9158294796943665
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.37it/s, pos cos_sim: 0.987
neg cos_sim: 0.983]

Average cos sim (pos, neg): 0.983893, 0.979352
Epoch: 6 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 54.09it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:49<00:00, 1.32s/it, loss: 1.484]

Adjusting learning rate of group 0 to 1.8875e-07.
Average epoch loss: 1.513662301650768
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.37it/s, pos cos_sim: 0.987
neg cos_sim: 0.982]

Average cos sim (pos, neg): 0.984632, 0.978742
Epoch: 7 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 54.10it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:49<00:00, 1.32s/it, loss: 1.414]

Adjusting learning rate of group 0 to 1.8875e-07.
Average epoch loss: 1.3313449514353006
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.38it/s, pos cos_sim: 0.987
neg cos_sim: 0.981]

Average cos sim (pos, neg): 0.985008, 0.977833
Epoch: 8 of 10

Generate siamese dataset

claims: 100%| | 1228/1228 [00:22<00:00, 54.02it/s]

Generated data n=19860

train batches: 100%| | 311/311 [06:48<00:00, 1.31s/it, loss: 1.801]

Adjusting learning rate of group 0 to 1.8875e-07.
Average epoch loss: 1.2004387966593746
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%| | 32/32 [00:07<00:00, 4.37it/s, pos cos_sim: 0.986
neg cos_sim: 0.981]

Average cos sim (pos, neg): 0.984882, 0.977008
Epoch: 9 of 10

```
Generate siamese dataset

claims: 100%|        | 1228/1228 [00:22<00:00, 53.71it/s]

Generated data n=19860

train batches: 100%|        | 311/311 [06:49<00:00,  1.32s/it, loss: 1.875]

Adjusting learning rate of group 0 to 1.8875e-07.
Average epoch loss: 1.063198039242309
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03a
_bert_base_triplet_2023_05_04_01_42.pth

dev batches: 100%|        | 32/32 [00:07<00:00,  4.40it/s, pos cos_sim: 0.986
neg cos_sim: 0.980]

Average cos sim (pos, neg): 0.984960, 0.976186
Done!
```

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

```python
# 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,  4.40it/s, pos cos_sim: 0.986
neg cos_sim: 0.980]
```

```
[ ]: 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): 0.984960, 0.976186
```

```
[ ]: sentence_rankings[:50]
```

```
[ ]: [SentenceSimilarity(label=1, score=0.99715555, 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=0.9907407, 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=0.9814228, 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=0.9634737, 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=0.9902761, 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=0.98679817, 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=0.9867811, 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=0.9759532, 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=0.9895188, 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=0.9905056, 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=0.9858325, 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=0.9833054, 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=0.992003, 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]agricultural operations, excluding land use changes, produce approximately 13 per cent of anthropogenic global GHG emissions.'),
 SentenceSimilarity(label=1, score=0.9874858, 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=0.98083556, 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=0.9953047, 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 CO 2 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=0.9931502, 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=0.9754391, 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=0.958438, 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=0.98822165, 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=0.9837831, 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=0.98180556, 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=0.9868192, 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=0.9880688, 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=0.98693156, 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=0.9821743, 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=0.9844698, 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=0.9951848, 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=0.9983725, 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=0.9901616, 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=0.9311247, 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=0.98707074, 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=0.98878956, 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=0.9844391, 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=0.95794696, 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=0.98791724, 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=0.98258346, 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=0.99180317, 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=0.9901115, 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=0.9549488, 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=0.986797, 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=0.9950853, 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=0.992875, 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=0.99060845, 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=0.9798428, 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=0.96060336, 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=0.9919453, 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=0.9895648, 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=0.98760694, 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=0.9775828, 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.')]
```

```python
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=0.883747, claim_text='That's because as
Antarctica's mass shrinks, the ice sheet's gravitational pull on the ocean
relaxes somewhat, and the seas travel back across the globe to pile up far away
- with U.S. coasts being one prime destination."', evidence_text='With the loss
of mass, the gravitational pull becomes less and local water levels might
drop.'),
 SentenceSimilarity(label=1, score=0.91259325, 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=0.9210543, 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=0.9303059, 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=0.9412791, claim_text='Concentrated in the
atmosphere, these gases do not allow the warmth of the sun's rays reflected by
the earth to be dispersed in space.', evidence_text='Ozone acts as a greenhouse
gas, absorbing some of the infrared energy emitted by the earth.'),
 SentenceSimilarity(label=1, score=0.9444597, 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='They measure radiances
in various wavelength bands.'),
 SentenceSimilarity(label=1, score=0.94595, 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=0.94782454, 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='Many feasible scenarios can be constructed to account for evidence.'),
 SentenceSimilarity(label=1, score=0.9482251, 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=0.9488129, 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=0.94951516, 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=0.95052344, 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=0.9511937, 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=0.95382774, claim_text='[…] Constant 24-7 media coverage of every significant storm worldwide just makes it seem that way."', evidence_text='The severe and widespread damage the storm caused in the United States, as well as its unusual merger with a frontal system, resulted in the nicknaming of the hurricane "Superstorm Sandy" by the media, public officials, and several organizations, including U.S. government agencies.'),
 SentenceSimilarity(label=1, score=0.9538534, 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='There is a large potential for future reductions in emissions by a combination of activities, including energy conservation and increased energy efficiency; the use of low-carbon energy technologies, such as renewable energy, nuclear energy, and carbon capture and storage; decarbonizing buildings and transport; and enhancing carbon sinks through, for example, reforestation and preventing deforestation.'),
 SentenceSimilarity(label=1, score=0.95403457, claim_text='Jim Hansen had several possible scenarios; his mid-level scenario B was right.',

evidence_text='In 2000 Hansen authored a paper called "Global warming in the twenty-first century: an alternative scenario" in which he presented a more optimistic way of dealing with global warming, focusing on non-CO2 gases and black carbon in the short run, giving more time to make reductions in fossil fuel emissions.'),
 SentenceSimilarity(label=1, score=0.9548111, 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=0.9549488, 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=0.9550726, 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=0.9559422, 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='Therefore, scenarios that are not justified by other evidence or are created to account for the historical evidence have not been universally accepted.'),
 SentenceSimilarity(label=1, score=0.95864475, 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=0.959322, 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=0.95995253, 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='To obtain milk from dairy cattle, cows are made pregnant to induce lactation; they are kept lactating for three to seven years, then slaughtered.'),
 SentenceSimilarity(label=1, score=0.9611503, claim_text='Greenpeace didn't save the whales, switching from whale oil to petroleum and palm oil did', evidence_text='Greenpeace is also critical of extracting petroleum from oil sands and has used direct action to block operations at the Athabasca oil sands in Canada.'),
 SentenceSimilarity(label=1, score=0.96162176, 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=0.96253055, 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=0.9636339, claim_text='Any reasonable person can recognize both positives and negatives among the policy proposals of both Tories and Labour.', evidence_text='The Liberal Democrats, the Greens, the SNP and Labour all support a ban on fracking, whilst the Conservatives propose approving fracking on a case-by-case basis.'),
 SentenceSimilarity(label=1, score=0.9640618, claim_text='as the pathbreaking work by Rosamond Naylor and David Battisti has shown, the tropics are already too hot to efficiently grow grain, and those places where grain is produced today are already at optimal growing temperature – which means even a small warming will push them down the slope of declining productivity.', evidence_text='Rice becomes sterile if exposed to temperatures above 35 degrees for more than one hour during flowering and consequently produces no grain.'),
 SentenceSimilarity(label=1, score=0.96442115, claim_text='"The solubility of carbon dioxide in water is listed in the Handbook of Chemistry and Physics as a declining function of temperature. …', evidence_text='In liquid water at high temperatures, (e.g.'),
 SentenceSimilarity(label=1, score=0.9646927, 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=0.96504986, claim_text='Last year's warmth was manifested across the planet, from the warm tropical ocean waters off the coast of northeastern Australia, where the Great Barrier Reef experienced its worst coral bleaching event on record and large scale coral death, to the Arctic, where sea ice hit regular monthly record lows and overall temperatures were also the warmest on record, at least from January through September 2016.', evidence_text='A March 2016 report stated that coral bleaching was more widespread than previously thought, seriously affecting the northern parts of the reef as a result of warming ocean temperatures.'),
 SentenceSimilarity(label=1, score=0.9657351, claim_text='"In their award winning book, 'Taken By Storm' (2007), Canadian researchers Christopher Essex and Ross McKitrick explain: 'Temperature is not an amount of something [like height or weight].', evidence_text='It has the symbol K. For everyday applications, it is often convenient to use the Celsius scale, in which 0\xa0°C corresponds very closely to the freezing point of water and 100\xa0°C is its boiling point at sea level.'),
 SentenceSimilarity(label=1, score=0.96651983, claim_text='More importantly, the

OISM list only contains 39 scientists who specialise in climate science.',
evidence_text='Ongoing experiments are conducted by more than 4,000 scientists
from many nations.'),
 SentenceSimilarity(label=1, score=0.96686316, 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=0.9668771, claim_text='Heat is continuing to
build up in the subsurface ocean.', evidence_text='It plays an important role in
supplying heat to the polar regions, and thus in sea ice regulation.'),
 SentenceSimilarity(label=1, score=0.96690756, 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='Cosmic-temperature measurements The VLT has detected, for the
first time, carbon-monoxide molecules in a galaxy located almost 11\xa0billion
light-years away.'),
 SentenceSimilarity(label=1, score=0.9673245, claim_text="Al Gore's book is
quite accurate, and far more accurate than contrarian books.", evidence_text='In
a 2007 court case, a British judge said that while he had "no doubt …the film
was broadly accurate" and its "four main scientific hypotheses …are supported
by a vast quantity of research", he upheld nine of a "long schedule" of alleged
errors presented to the court.'),
 SentenceSimilarity(label=1, score=0.96765816, 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=0.9686582, claim_text='Concentrated in the
atmosphere, these gases do not allow the warmth of the sun's rays reflected by
the earth to be dispersed in space.', evidence_text='This last phenomenon is
known as the greenhouse effect: trace molecules within the atmosphere serve to
capture thermal energy emitted from the ground, thereby raising the average
temperature.'),
 SentenceSimilarity(label=1, score=0.9695087, 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='"Ireland to Plant 440 Million
Trees in 20 Years to Fight Climate Change".'),
 SentenceSimilarity(label=1, score=0.9699521, claim_text='‘Next year or the year
after, the Arctic will be free of ice’', evidence_text='"US Navy predicts summer
ice free Arctic by 2016".'),
 SentenceSimilarity(label=1, score=0.9702539, 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=0.9704229, 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 Intergovernmental Panel on Climate Change said the likelihood was 90 percent to 99 percent that emissions of heat-trapping greenhouse gases like carbon dioxide, spewed from tailpipes and smokestacks, were the dominant cause of the observed warming of the last 50 years.'),
 SentenceSimilarity(label=1, score=0.97060615, 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='On the contrary, there are reasons to believe that this is altogether out of the question."'),
 SentenceSimilarity(label=1, score=0.9707868, 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='"Storm Harvey: impacts likely worsened due to global warming".'),
 SentenceSimilarity(label=1, score=0.9708744, 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='Carnivores require meat in their diet for survival, but carnists choose to eat meat based on their beliefs.'),
 SentenceSimilarity(label=1, score=0.9708824, claim_text='"In their award winning book, 'Taken By Storm' (2007), Canadian researchers Christopher Essex and Ross McKitrick explain: 'Temperature is not an amount of something [like height or weight].', evidence_text='The most common scales are the Celsius scale (formerly called centigrade), denoted °C, the Fahrenheit scale (denoted °F), and the Kelvin scale (denoted K), the latter of which is predominantly used for scientific purposes by conventions of the International System of Units (SI).'),
 SentenceSimilarity(label=1, score=0.97111875, 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=0.9711963, claim_text='The rate of renewable energy installations in the EU in 2018 was less than half the maximum level achieved in 2010.', evidence_text='In January 2014, the EU agreed to a 40% emissions reduction by 2030, compared to 1990 levels, and a 27% renewable energy target, which is expected to provide 70,000 full-time jobs and cut €33bn in fossil fuel imports.'),
 SentenceSimilarity(label=1, score=0.97135925, claim_text='Through its impacts on the climate, CO2 presents a danger to public health and welfare, and thus qualifies as an air pollutant', evidence_text='An air pollutant is a material in

the air that can have adverse effects on humans and the ecosystem.')]

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

```
[ ]: [SentenceSimilarity(label=-1, score=0.99675167, 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='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=0.9966569, 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 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=0.99660766, 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='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=0.99654603, claim_text='The majority of peer
     reviewed research at the time predicted warming due to increasing CO2.',
     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=0.99648064, claim_text='In many other cases,
     though – hurricanes, for example – the linkage to global warming for particular
     trends is uncertain or disputed.', 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=0.9964528, 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='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=0.99636567, claim_text='Latest IPCC Reports
     (AR5) have shown global mean temperature forecasts from the 2005 IPCC report
     exceeded actual readings.', 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=0.9963034, 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=0.996073, claim_text='The amount of carbon dioxide absorbed by the upper layer of the oceans is increasing by about 2 billion tons per year.', 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=0.99580055, 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=0.9957638, claim_text='Scientists studying Antarctica sea ice warn a rise in accumulation could spark the next ice age.', 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=0.9957619, 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='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=0.9957467, 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='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=0.9956937, claim_text='Scientists studying Antarctica sea ice warn a rise in accumulation could spark the next ice age.', 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=0.99566877, claim_text='Once natural influences, in particular the impact of El Niño and La Niña, are removed from the recent termperature 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=0.9955871, claim_text='The most recent IPCC report lays out a future if we limit global heating to 1.5°C instead of the

Paris Agreement's 2°C.', 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=0.9955702, 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=0.9955288, claim_text='In many other cases, though - hurricanes, for example - the linkage to global warming for particular trends is uncertain or disputed.', 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=0.99547434, 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='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=0.99539125, claim_text="There are a myriad of other radiative forcings that affect the planet's energy imbalance.", 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=0.99530435, 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=0.99526656, 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=0.9952383, claim_text='IPCC overestimate temperature rise.', 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=0.9951836, claim_text='the models predicted seven times as much warming as has been observed', 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=0.99505764, claim_text='IPCC graph showing accelerating trends is misleading', evidence_text='"How the oceans absorb carbon dioxide is critical for predicting climate change".'),

SentenceSimilarity(label=-1, score=0.99500084, claim_text='Scientists studying Antarctica sea ice warn a rise in accumulation could spark the next ice age.', 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=0.9949881, claim_text='Global Warming history completely coincides with the history of artificial satellites and the use of microwave frequencies from outer space.', 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=0.99494445, 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='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=0.99489444, claim_text='after a natural orbitally driven warming, atmospheric carbon dioxide content increases 800 years later', 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=0.99489343, claim_text='While there are isolated cases of growing glaciers, the overwhelming trend in glaciers worldwide is 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=0.9948405, 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=0.99483585, claim_text='The long term trend from albedo is that of cooling.', evidence_text='"How the oceans absorb carbon dioxide is critical for predicting climate change".'),
SentenceSimilarity(label=-1, score=0.99481475, 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=0.9948093, claim_text="Global warming' and 'climate change' mean different things and have both been used for decades.", 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=0.99477744, claim_text='IPCC report warning last week the world is "nowhere near on track" to meet its Paris commitments', 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=0.99460363, 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=0.9945767, claim_text='Once natural influences, in particular the impact of El Niño and La Niña, are removed from the recent termperature record, there is no evidence of a significant change in the human contribution to climate change.', 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=0.99439865, 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=0.99437165, claim_text='While the greenhouse effect is a natural occurence, too much warming has severe negative impacts on agriculture, health and environment.', 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=0.99432296, claim_text='IPCC graph showing accelerating trends is misleading', evidence_text='Scientists Reach 100% Consensus on Anthropogenic Global Warming.'),
 SentenceSimilarity(label=-1, score=0.99430096, claim_text='Venus very likely underwent a runaway or 'moist' greenhouse phase earlier in its history, and today is kept hot by a dense CO2 atmosphere.', 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=0.99425405, 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=0.9941981, claim_text='The heatwave we now have in Europe is not something that was expected with just 1C of warming', evidence_text='"How the oceans absorb carbon dioxide is critical for predicting

climate change".'),
 SentenceSimilarity(label=-1, score=0.99409544, 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=0.994085, claim_text='the models predicted seven times as much warming as has been observed', evidence_text='"How the oceans absorb carbon dioxide is critical for predicting climate change".'),
 SentenceSimilarity(label=-1, score=0.99401593, 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='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=0.9940116, claim_text='"Twentieth century global warming did not start until 1910.', 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=0.9940113, claim_text='[…] Constant 24-7 media coverage of every significant storm worldwide just makes it seem that way."', 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=0.99397796, 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=0.993898, claim_text='the models predicted seven times as much warming as has been observed', 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.')]