model_03_er_siamese_bert_base

May 2, 2023

1 Model 03

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 Linear, Module, CrossEntropyLoss, Dropout
from transformers import BertModel, BertTokenizer
from torch.optim import Adam
from torch.optim.lr_scheduler import LinearLR
from torch.nn.functional import relu, softmax
from torcheval.metrics import BinaryAccuracy, BinaryF1Score, BinaryRecall
```

```
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
```

[ ]:
```
class SiameseDataset(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(SiameseDataset, 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
        ):
            # Check if we have evidences supplied, this will inform
            # whether this is for training
            is_training = "evidences" in claim.keys()
```

```
            pos_evidence_ids = set()

            # Get positive samples from evidences with label=1
            if is_training:
                pos_evidence_ids.update(claim["evidences"])

                for evidence_id in pos_evidence_ids:
                    data.append(ClaimEvidencePair(
                        claim_id=claim_id,
                        evidence_id=evidence_id,
                        label=1
                    ))

            # Get negative samples from pre-retrieved evidences
            # for each claim with label=0
            retrieved_evidence_ids = self.claims_shortlist.get(claim_id, [])
            if len(retrieved_evidence_ids) > 0:
                retrieved_neg_evidence_ids = random.sample(
                    population=retrieved_evidence_ids,
                    k=min(self.n_neg_shortlist, len(retrieved_evidence_ids))
                )

                # Generate claim and shortlisted negative evidence pairs
                for evidence_id in retrieved_neg_evidence_ids:
                    data.append(ClaimEvidencePair(
                        claim_id=claim_id,
                        evidence_id=evidence_id,
                        label=0
                    ))

            # Get negative samples from shortlisted evidences list with label=0
            if len(self.evidence_shortlist) > 0:
                shortlist_neg_evidence_ids = random.sample(
                    population=self.evidence_shortlist,
                    k=min(self.n_neg_general, len(self.evidence_shortlist))
                )

                # Generate claim and shortlisted negative evidence pairs
                for evidence_id in shortlist_neg_evidence_ids:
                    data.append(ClaimEvidencePair(
                        claim_id=claim_id,
                        evidence_id=evidence_id,
                        label=0
                    ))

            continue
```

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

        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)
```

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

# train_data = SiameseDataset(
#     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=100,
#     n_neg_general=100
# )
```

1.3 Build model

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

    def __init__(
            self,
            pretrained_name:str,
            device,
            **kwargs
```

```
        ) -> None:
        super(SiameseClassifierBert, 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)

        # Classification layers
        self.linear1 = Linear(2304, 1024, bias=True, device=device)
        self.linear2 = Linear(1024, 512, bias=True, device=device)
        self.linear3 = Linear(512, 2, bias=True, device=device)
        self.relu = relu
        self.softmax = softmax
        self.dropout_in = Dropout(p=0.2)
        self.dropout_out = Dropout(p=0.5)

        # print(self.tokenizer)
        # print(self.bert)
        # print(self.linear1)
        # print(self.linear2)
        # print(self.activation)
        # print(self.softmax)
        return

    def forward(self, claim_texts, evidence_texts) -> Tuple[torch.Tensor]:

        # Run the tokenizer
        t_kwargs = {
            "return_tensors": "pt",
            "padding": True,
            "truncation": True,
            "max_length": 100,
            "add_special_tokens":True
        }
        claim_x = self.tokenizer(claim_texts, **t_kwargs)
        evidence_x = self.tokenizer(evidence_texts, **t_kwargs)

        claim_x = claim_x["input_ids"].to(device=self.device)
        evidence_x = evidence_x["input_ids"].to(device=self.device)

        # Run Bert
        claim_x = self.bert(claim_x, return_dict=True).pooler_output
        evidence_x = self.bert(evidence_x, return_dict=True).pooler_output
```

```
        # dim=768

        # Concatenate the two embeddings
        x = torch.cat((claim_x, evidence_x, claim_x - evidence_x), dim=1)
        # dim=2304

        # Run classification layers
        x = self.dropout_in(x)
        x = self.linear1(x)
        x = self.relu(x)
        x = self.dropout_out(x)
        x = self.linear2(x)
        x = self.relu(x)
        x = self.dropout_out(x)
        x = self.linear3(x)

        # Create the predictions
        y = self.softmax(x, dim=-1)

        return (y, claim_x, evidence_x)
```

1.4 Training and evaluation loop

[ ]:
```
model = SiameseClassifierBert(
    pretrained_name="bert-base-uncased",
    device=TORCH_DEVICE
)
```

```
Some weights of the model checkpoint at bert-base-uncased were not used when
initializing BertModel: ['cls.predictions.transform.dense.bias',
'cls.predictions.bias', 'cls.seq_relationship.bias',
'cls.predictions.transform.LayerNorm.bias',
'cls.predictions.transform.LayerNorm.weight', 'cls.predictions.decoder.weight',
'cls.seq_relationship.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 = CrossEntropyLoss()
optimizer = Adam(
    params=model.parameters(),
    lr=0.000002
) #! Hyperparams
```

[ ]:
```
run_time = datetime.now().strftime('%Y_%m_%d_%H_%M')
MODEL_NAME = f"model_03_run_01_base_{run_time}.pth"
N_EPOCHS = 100
BATCH_SIZE = 64
```

[ ]:
```
dev_data = SiameseDataset(
    claims_paths=[DATA_PATH.with_name("dev-claims.json")],
    claims_shortlist_paths=[NER_PATH.with_name("dev_claim_evidence_retrieved.
 ↪json")],
    evidence_shortlists=[NER_PATH.
 ↪with_name("shortlist_dev_claim_evidence_retrieved.json")],
    evidence_path=DATA_PATH.with_name("evidence.json"),
    device=TORCH_DEVICE,
    n_neg_shortlist=3,
    n_neg_general=3
)

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

```
Generate siamese dataset

claims: 100%|      | 154/154 [00:00<00:00, 478.61it/s]

Generated data n=1404
```

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

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

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

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

    # Run training
    model.train()

    train_data = SiameseDataset(
        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:
        claim_texts, evidence_texts, labels = batch

        # Reset optimizer
        optimizer.zero_grad()

        # Forward + loss
        predictions, *_ = model(claim_texts, evidence_texts)
        loss = loss_fn(predictions, labels)

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

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

    # Run evaluation
    model.eval()

    dev_batches = tqdm(dev_dataloader, desc="dev batches")
    dev_acc = []
    dev_f1 = []
    dev_rec = []
    for batch in dev_batches:
        claim_texts, evidence_texts, labels = batch

        # Forward
        predictions, *_ = model(claim_texts, evidence_texts)

        # Prediction
        _, predicted = torch.max(predictions, dim=-1)

        # Metrics
        metric_accuracy.update(predicted.cpu(), labels.cpu())
        metric_f1.update(predicted.cpu(), labels.cpu())
        metric_recall.update(predicted.cpu(), labels.cpu())

        acc = metric_accuracy.compute()
        f1 = metric_f1.compute()
        rec = metric_recall.compute()

        dev_acc.append(acc)
        dev_f1.append(f1)
        dev_rec.append(rec)

        dev_batches.postfix = \
            f" acc: {acc:.3f}" \
            + f" f1: {f1:.3f}" \
```

```
            + f" rec: {rec:.3f}"

        continue

    val_acc = np.mean(dev_acc)
    val_f1 = np.mean(dev_f1)
    val_rec = np.mean(dev_rec)

    print(f"Epoch accuracy on dev: {val_acc:.3f}")
    print(f"Epoch f1 on dev: {val_f1:.3f}")
    print(f"Epoch recall on dev: {val_rec:.3f}\n")

print("Done!")
```

```
Adjusting learning rate of group 0 to 2.0000e-07.
Epoch: 0 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 181.77it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:47<00:00,  1.13it/s, loss: 2.097]

Adjusting learning rate of group 0 to 3.8000e-07.
Average epoch loss: 2.081911284076281
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|      | 22/22 [00:07<00:00,  2.98it/s,  acc: 0.591 f1:
0.282 rec: 0.282]

Epoch accuracy on dev: 0.594
Epoch f1 on dev: 0.290
Epoch recall on dev: 0.290

Epoch: 1 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:07<00:00, 174.75it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.14it/s, loss: 2.086]

Adjusting learning rate of group 0 to 5.6000e-07.
Average epoch loss: 2.078024445486463
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 2 of 100
```

```
Generate siamese dataset

claims: 100%|      | 1228/1228 [00:07<00:00, 166.58it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:46<00:00,  1.13it/s, loss: 2.097]

Adjusting learning rate of group 0 to 7.4000e-07.
Average epoch loss: 2.0727585258562704
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 3 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 178.99it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:46<00:00,  1.14it/s, loss: 2.035]

Adjusting learning rate of group 0 to 9.2000e-07.
Average epoch loss: 2.0598962341458344
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 4 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 178.16it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:46<00:00,  1.14it/s, loss: 1.956]

Adjusting learning rate of group 0 to 1.1000e-06.
Average epoch loss: 1.946168662595355
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 5 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 179.83it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.665]

Adjusting learning rate of group 0 to 1.2800e-06.
Average epoch loss: 1.7010489135734306
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
```

```
dev batches: 100%|     | 22/22 [00:05<00:00,  3.74it/s,  acc: 0.675 f1:
0.557 rec: 0.557]

Epoch accuracy on dev: 0.644
Epoch f1 on dev: 0.471
Epoch recall on dev: 0.471

Epoch: 6 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 179.06it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:46<00:00,  1.14it/s, loss: 1.475]

Adjusting learning rate of group 0 to 1.4600e-06.
Average epoch loss: 1.5637229316490742
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 7 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:07<00:00, 174.59it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.14it/s, loss: 1.555]

Adjusting learning rate of group 0 to 1.6400e-06.
Average epoch loss: 1.4980726953872965
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 8 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 178.58it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.14it/s, loss: 1.083]

Adjusting learning rate of group 0 to 1.8200e-06.
Average epoch loss: 1.467641536361915
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 9 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 179.73it/s]
```

```
Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.930]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.442951360517297
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 10 of 100


Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 179.65it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.14it/s, loss: 1.530]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.4167648615423314
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|      | 22/22 [00:05<00:00,  3.74it/s,  acc: 0.718 f1:
0.630 rec: 0.630]

Epoch accuracy on dev: 0.701
Epoch f1 on dev: 0.602
Epoch recall on dev: 0.602


Epoch: 11 of 100


Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 177.11it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.14it/s, loss: 1.154]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.3864337740358241
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 12 of 100


Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.23it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.078]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.3526843131081132
```

```
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 13 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 188.99it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 1.447]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.3503698814998975
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 14 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.45it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 1.177]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.3346617593745556
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 15 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.53it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 1.564]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.309252381324768
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|     | 22/22 [00:05<00:00,  3.76it/s,  acc: 0.745 f1:
0.671 rec: 0.671]

Epoch accuracy on dev: 0.733
Epoch f1 on dev: 0.654
Epoch recall on dev: 0.654

Epoch: 16 of 100

Generate siamese dataset
```

```
claims: 100%|      | 1228/1228 [00:06<00:00, 187.69it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 0.955]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.3000709375074087
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 17 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 188.23it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.326]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2864724252342192
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 18 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.51it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.415]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2818637184860293
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 19 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 188.30it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.15it/s, loss: 1.134]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.27057289461459
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 20 of 100

Generate siamese dataset
```

```
claims: 100%|      | 1228/1228 [00:06<00:00, 187.35it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.387]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2467697747975341
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|      | 22/22 [00:05<00:00,  3.74it/s,  acc: 0.763 f1:
0.696 rec: 0.696]

Epoch accuracy on dev: 0.755
Epoch f1 on dev: 0.686
Epoch recall on dev: 0.686

Epoch: 21 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.76it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.464]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2537854759653737
Epoch: 22 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 186.28it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.15it/s, loss: 1.434]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2495689924098243
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 23 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.94it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.470]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2454679480268935
```

```
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 24 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.52it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 0.950]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2354927173823365
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 25 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.89it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.15it/s, loss: 0.946]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2366248179565777

dev batches: 100%|     | 22/22 [00:05<00:00,  3.72it/s,  acc: 0.775 f1:
0.710 rec: 0.710]

Epoch accuracy on dev: 0.769
Epoch f1 on dev: 0.704
Epoch recall on dev: 0.704

Epoch: 26 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 186.28it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 1.413]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.226368307082121
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 27 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.51it/s]
```

```
Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.232]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2220329226048525
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 28 of 100


Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 188.45it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.15it/s, loss: 1.173]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.217984320949917
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 29 of 100


Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.75it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.721]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.222490115845499
Epoch: 30 of 100


Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 182.64it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.302]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2196419874498667
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|      | 22/22 [00:05<00:00,  3.75it/s,  acc: 0.783 f1:
0.720 rec: 0.720]

Epoch accuracy on dev: 0.779
Epoch f1 on dev: 0.716
Epoch recall on dev: 0.716
```

```
Epoch: 31 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.84it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.630]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2223843252363285
Epoch: 32 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.60it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 0.956]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2107447735534227
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 33 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.97it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.356]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1992103467302875
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 34 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 188.46it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.15it/s, loss: 1.517]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.198439207944003
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 35 of 100
```

```
Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.55it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 1.398]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.2036892648570794

dev batches: 100%|     | 22/22 [00:05<00:00,  3.77it/s,  acc: 0.789 f1:
0.727 rec: 0.727]

Epoch accuracy on dev: 0.786
Epoch f1 on dev: 0.724
Epoch recall on dev: 0.724

Epoch: 36 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 189.13it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.14it/s, loss: 1.173]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1940935416655107
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 37 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.93it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 1.296]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1804965537441663
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 38 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 188.33it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 1.402]
```

```
Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1883822098251218
Epoch: 39 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.56it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 1.403]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1933931023621362
Epoch: 40 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 186.33it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 0.941]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1876138293546092
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|     | 22/22 [00:05<00:00,  3.76it/s,  acc: 0.794 f1:
0.733 rec: 0.733]

Epoch accuracy on dev: 0.792
Epoch f1 on dev: 0.730
Epoch recall on dev: 0.730

Epoch: 41 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 186.84it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.15it/s, loss: 1.164]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1680641772825855
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 42 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.06it/s]
```

```
Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.548]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1880789391758029
Epoch: 43 of 100


Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.55it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.401]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.177235478950926
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 44 of 100


Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 188.73it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 0.941]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1698007184611865
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 45 of 100


Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 188.18it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.010]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.162754298980571
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|      | 22/22 [00:05<00:00,  3.75it/s,  acc: 0.799 f1:
0.737 rec: 0.737]

Epoch accuracy on dev: 0.797
Epoch f1 on dev: 0.735
Epoch recall on dev: 0.735
```

```
Epoch: 46 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.72it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.226]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1654598757255177
Epoch: 47 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 186.75it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.177]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1680036731495345
Epoch: 48 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.15it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.411]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.159905521091351
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 49 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 186.41it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 0.943]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1531196327741482
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 50 of 100

Generate siamese dataset
```

```
claims: 100%|      | 1228/1228 [00:06<00:00, 187.71it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 0.985]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1680524006855388

dev batches: 100%|      | 22/22 [00:05<00:00,  3.72it/s,  acc: 0.803 f1:
0.742 rec: 0.742]

Epoch accuracy on dev: 0.801
Epoch f1 on dev: 0.740
Epoch recall on dev: 0.740

Epoch: 51 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.97it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.14it/s, loss: 1.171]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1443446212071033
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 52 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 188.51it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.15it/s, loss: 1.922]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1654003728027187
Epoch: 53 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 182.95it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.172]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1565275662694097
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
```

```
Epoch: 54 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.90it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.178]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1552124163828605
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 55 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.37it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.081]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.158671263820869

dev batches: 100%|      | 22/22 [00:05<00:00,  3.71it/s,  acc: 0.806 f1:
0.745 rec: 0.745]

Epoch accuracy on dev: 0.805
Epoch f1 on dev: 0.743
Epoch recall on dev: 0.743

Epoch: 56 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 185.86it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:43<00:00,  1.17it/s, loss: 1.104]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1591322018095285
Epoch: 57 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.45it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.201]
```

```
Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1570096796701763
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 58 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.89it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 0.940]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1511653590300852
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 59 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.68it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.15it/s, loss: 1.567]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1510806994989884
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 60 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.54it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.029]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1476330582267982
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|      | 22/22 [00:05<00:00,  3.73it/s,  acc: 0.809 f1:
0.748 rec: 0.748]

Epoch accuracy on dev: 0.808
Epoch f1 on dev: 0.746
Epoch recall on dev: 0.746

Epoch: 61 of 100
```

```
Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.66it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 1.163]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1593439322857817
Epoch: 62 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.00it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 1.172]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1539781724125886
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 63 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 188.37it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.15it/s, loss: 0.940]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1274150425244953
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 64 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 188.01it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:43<00:00,  1.17it/s, loss: 1.034]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1437808245666756
Epoch: 65 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 183.72it/s]
```

```
Generated data n=7693

train batches: 100%|      | 121/121 [01:43<00:00,  1.16it/s, loss: 1.211]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1362658276045619
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|      | 22/22 [00:05<00:00,  3.75it/s,  acc: 0.812 f1:
0.749 rec: 0.749]

Epoch accuracy on dev: 0.810
Epoch f1 on dev: 0.749
Epoch recall on dev: 0.749

Epoch: 66 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 188.61it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.051]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1347789498400096
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 67 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 186.46it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.172]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1284156557449625
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 68 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.81it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.171]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1376145147095043
```

```
Epoch: 69 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 186.48it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.573]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1308643278504207
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 70 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.06it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.114]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1369210720554856

dev batches: 100%|      | 22/22 [00:05<00:00,  3.73it/s,  acc: 0.814 f1:
0.751 rec: 0.751]

Epoch accuracy on dev: 0.813
Epoch f1 on dev: 0.750
Epoch recall on dev: 0.750

Epoch: 71 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 186.09it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.189]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.120794879996087
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 72 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.46it/s]

Generated data n=7693
```

```
train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 1.102]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1227812855696875
Epoch: 73 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 186.29it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:46<00:00,  1.14it/s, loss: 1.171]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1513384531352147
Epoch: 74 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.91it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 1.401]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1274598435429501
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 75 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.40it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 1.173]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.144767465916547

dev batches: 100%|     | 22/22 [00:05<00:00,  3.75it/s,  acc: 0.815 f1:
0.752 rec: 0.752]

Epoch accuracy on dev: 0.814
Epoch f1 on dev: 0.751
Epoch recall on dev: 0.751

Epoch: 76 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.20it/s]
```

```
Generated data n=7693

train batches: 100%|      | 121/121 [01:43<00:00,  1.16it/s, loss: 0.940]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1135745558364332
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 77 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.32it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.15it/s, loss: 1.000]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.135797239039555
Epoch: 78 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 185.60it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.192]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.135993244973096
Epoch: 79 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.22it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:43<00:00,  1.16it/s, loss: 0.940]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1274315244402766
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 80 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 187.91it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.401]
```

```
Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1272837635406778
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|     | 22/22 [00:05<00:00,  3.74it/s,  acc: 0.817 f1:
0.753 rec: 0.753]

Epoch accuracy on dev: 0.816
Epoch f1 on dev: 0.752
Epoch recall on dev: 0.752

Epoch: 81 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.31it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 1.411]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1266227541876233
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 82 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.62it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.15it/s, loss: 0.941]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1253405370988137
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 83 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.87it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 0.951]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1048035232488773
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 84 of 100
```

```
Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.55it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 1.645]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1214817198347453
Epoch: 85 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.18it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 1.087]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1271609183677957

dev batches: 100%|     | 22/22 [00:05<00:00,  3.74it/s,  acc: 0.818 f1:
0.753 rec: 0.753]

Epoch accuracy on dev: 0.817
Epoch f1 on dev: 0.753
Epoch recall on dev: 0.753

Epoch: 86 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 186.27it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 1.399]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1270839389690683
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 87 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.60it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:43<00:00,  1.17it/s, loss: 1.172]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1142758699980648
```

```
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 88 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 183.56it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 0.990]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1208055164695772
Epoch: 89 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.31it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:45<00:00,  1.15it/s, loss: 1.410]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.115254128521139
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 90 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 187.77it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 0.943]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.105979498999178
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth

dev batches: 100%|     | 22/22 [00:05<00:00,  3.73it/s,  acc: 0.819 f1:
0.754 rec: 0.754]

Epoch accuracy on dev: 0.819
Epoch f1 on dev: 0.754
Epoch recall on dev: 0.754

Epoch: 91 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 186.38it/s]
```

```
Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 0.951]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1174351206003141
Epoch: 92 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 186.17it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.15it/s, loss: 1.171]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1043047498572955
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 93 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 186.99it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:44<00:00,  1.16it/s, loss: 1.401]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.117117817490554
Epoch: 94 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 185.82it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.170]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1013108171707342
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 95 of 100

Generate siamese dataset

claims: 100%|      | 1228/1228 [00:06<00:00, 188.18it/s]

Generated data n=7693

train batches: 100%|      | 121/121 [01:45<00:00,  1.15it/s, loss: 1.403]
```

```
Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1266289121848492

dev batches: 100%|     | 22/22 [00:05<00:00,  3.74it/s,  acc: 0.820 f1:
0.755 rec: 0.755]

Epoch accuracy on dev: 0.820
Epoch f1 on dev: 0.755
Epoch recall on dev: 0.755

Epoch: 96 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 186.60it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.15it/s, loss: 1.401]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1069179315212345
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Epoch: 97 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 186.74it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 0.983]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.1115300128282595
Epoch: 98 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 185.98it/s]

Generated data n=7693

train batches: 100%|     | 121/121 [01:44<00:00,  1.16it/s, loss: 0.949]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.116132346559162
Epoch: 99 of 100

Generate siamese dataset

claims: 100%|     | 1228/1228 [00:06<00:00, 186.64it/s]

Generated data n=7693
```

```
train batches: 100%|      | 121/121 [01:46<00:00,  1.14it/s, loss: 0.941]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.109876610523413
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_03_
run_01_base_2023_05_01_23_07.pth
Done!
```

model_03_run_01_base_2023_05_01_22_46: good candidate, LR = 0.00005, epoch loss = 1.272 after 9 epochs