model_04_label_nli_roberta_base_run_01

May 3, 2023

1 Model 04

Prediction of claim labels based on the matched evidence.

Ref: - Hugging face pre-trained models - Hugging face guide to fine-tuning - Hugging face guide to fine-tuning easy - SO Guide - roberta-large-mnli

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 Linear, Module, CrossEntropyLoss, Dropout
from transformers import RobertaModel, RobertaTokenizer
from torch.optim import Adam
from torch.optim.lr_scheduler import LinearLR
from torch.nn.functional import relu, softmax
from torcheval.metrics import MulticlassAccuracy, MulticlassF1Score

from src.torch_utils import get_torch_device
```

```python
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
class SiameseLabelDataset(Dataset):

    # Pre-trained model has:
    # label_map = {0: 'contradiction', 1: 'neutral', 2: 'entailment'}
    LABEL_MAP = {
        "REFUTES": [0],
        "NOT_ENOUGH_INFO": [1],
        "SUPPORTS": [2],
        "DISPUTED": [0, 2],
    }

    def __init__(
        self,
        claims_paths:List[Path],
        evidence_path:Path,
        device = None,
        verbose:bool=True
    ) -> None:
        super(SiameseLabelDataset, self).__init__()
        self.verbose = verbose
        self.device = device
```

```python
        # 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))

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

        # 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
        ):
            # Get evidence ids associated with each claim
            evidence_ids = claim["evidences"]

            # Get the claim label
            label = claim["claim_label"]

            # Encode the label
            label_encodings = self.LABEL_MAP.get(label, 99)

            if 99 in label_encodings:
                print(f"Invalid label {label} for {claim_id}")

            for label_encoding in label_encodings:

                # Assign the labels to each claim/evidence pairs
                for evidence_id in evidence_ids:
                    data.append(ClaimEvidencePair(
                        claim_id=claim_id,
                        evidence_id=evidence_id,
                        label=label_encoding
                    ))
```

3

```python
            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)
        # label = torch.unsqueeze(label, -1)

        # 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)
```

1.3 Build model

```python
class SiameseClassifierRoberta(Module):

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

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

        # Use a pretrained model
        self.bert = RobertaModel.from_pretrained(pretrained_name)
        self.bert.to(device=device)
```

```python
        # Classification layers
        self.linear1 = Linear(768, 256, bias=True, device=device)
        self.linear2 = Linear(256, 3, bias=True, device=device)
        self.relu = relu
        self.softmax = softmax
        self.dropout_in = Dropout(p=0.2)
        self.dropout_out = Dropout(p=0.5)
        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": 512,
            "add_special_tokens":True
        }
        # Use a cross-encoding pattern
        x = self.tokenizer(claim_texts, evidence_texts, **t_kwargs)
        x = x["input_ids"].to(device=self.device)
        x = self.bert(x, return_dict=True).pooler_output
        # dim=768

        # 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)

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

        return y
```

1.4 Training and evaluation loop

```python
CHECKPOINT_NAME = f"model_04_roberta_base_3_class.pth"
with open(MODEL_PATH.with_name(CHECKPOINT_NAME), mode="rb") as f:
    model = torch.load(f)
```

```python
# model = SiameseClassifierRoberta(
#     pretrained_name="roberta-base",
#     device=TORCH_DEVICE
```

```
# )
```

```
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_04_roberta_base_{run_time}.pth"
MODEL_NAME = CHECKPOINT_NAME
N_EPOCHS = 100
BATCH_SIZE = 24
```

```
train_data = SiameseLabelDataset(
    claims_paths=[DATA_PATH.with_name("train-claims.json")],
    evidence_path=DATA_PATH.with_name("evidence.json"),
    device=TORCH_DEVICE,
)

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

Generate siamese dataset

claims: 100%| | 1228/1228 [00:00<00:00, 566685.59it/s]

Generated data n=4514

```
dev_data = SiameseLabelDataset(
    claims_paths=[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
)
```

Generate siamese dataset

claims: 100%| | 154/154 [00:00<00:00, 594772.39it/s]

Generated data n=549

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

```python
metric_accuracy = MulticlassAccuracy()
metric_f1 = MulticlassF1Score()

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_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
```

```python
    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 = []
    for batch in dev_batches:
        claim_texts, evidence_texts, labels = batch

        # Forward
        predictions = model(claim_texts, evidence_texts)

        # Prediction
        predicted = torch.argmax(predictions, dim=1)

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

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

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

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

        continue

val_acc = np.mean(dev_acc)
val_f1 = np.mean(dev_f1)
```

```python
    print(f"Epoch accuracy on dev: {val_acc:.3f}")
    print(f"Epoch f1 on dev: {val_f1:.3f}")

print("Done!")
```

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

train batches: 100%| | 189/189 [01:17<00:00, 2.44it/s, loss: 2.110]

Adjusting learning rate of group 0 to 3.8000e-07.
Average epoch loss: 2.5900415816004316
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:04<00:00, 5.66it/s, acc: 0.472 f1:
0.472]

Epoch accuracy on dev: 0.471
Epoch f1 on dev: 0.471
Epoch: 1 of 100

train batches: 100%| | 189/189 [01:10<00:00, 2.69it/s, loss: 1.743]

Adjusting learning rate of group 0 to 5.6000e-07.
Average epoch loss: 2.574190516320486
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.79it/s, acc: 0.470 f1:
0.470]

Epoch accuracy on dev: 0.471
Epoch f1 on dev: 0.471
Epoch: 2 of 100

train batches: 100%| | 189/189 [01:07<00:00, 2.82it/s, loss: 1.760]

Adjusting learning rate of group 0 to 7.4000e-07.
Average epoch loss: 2.558910616806575
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.67it/s, acc: 0.472 f1:
0.472]

Epoch accuracy on dev: 0.471
Epoch f1 on dev: 0.471

```
Epoch: 3 of 100


train batches: 100%|      | 189/189 [01:07<00:00,  2.81it/s, loss: 4.395]

Adjusting learning rate of group 0 to 9.2000e-07.
Average epoch loss: 2.5703778068224588

dev batches: 100%|      | 23/23 [00:02<00:00,  9.55it/s,  acc: 0.476 f1:
0.476]

Epoch accuracy on dev: 0.474
Epoch f1 on dev: 0.474
Epoch: 4 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.85it/s, loss: 3.172]

Adjusting learning rate of group 0 to 1.1000e-06.
Average epoch loss: 2.540627516451336
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.79it/s,  acc: 0.482 f1:
0.482]

Epoch accuracy on dev: 0.479
Epoch f1 on dev: 0.479
Epoch: 5 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.83it/s, loss: 3.018]

Adjusting learning rate of group 0 to 1.2800e-06.
Average epoch loss: 2.530042542351617
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.12it/s,  acc: 0.485 f1:
0.485]

Epoch accuracy on dev: 0.484
Epoch f1 on dev: 0.484
Epoch: 6 of 100


train batches: 100%|      | 189/189 [01:07<00:00,  2.82it/s, loss: 1.738]

Adjusting learning rate of group 0 to 1.4600e-06.
Average epoch loss: 2.498065792378925
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.19it/s,  acc: 0.487 f1:
0.487]
```

Epoch accuracy on dev: 0.486
Epoch f1 on dev: 0.486
Epoch: 7 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.84it/s, loss: 2.770]

Adjusting learning rate of group 0 to 1.6400e-06.
Average epoch loss: 2.495964423058525
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.21it/s, acc: 0.488 f1:
0.488]

Epoch accuracy on dev: 0.488
Epoch f1 on dev: 0.488
Epoch: 8 of 100

train batches: 100%| | 189/189 [01:05<00:00, 2.88it/s, loss: 1.694]

Adjusting learning rate of group 0 to 1.8200e-06.
Average epoch loss: 2.4709754444303966
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.44it/s, acc: 0.487 f1:
0.487]

Epoch accuracy on dev: 0.488
Epoch f1 on dev: 0.488
Epoch: 9 of 100

train batches: 100%| | 189/189 [01:07<00:00, 2.82it/s, loss: 1.690]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.461363661856878
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.61it/s, acc: 0.489 f1:
0.489]

Epoch accuracy on dev: 0.487
Epoch f1 on dev: 0.487
Epoch: 10 of 100

train batches: 100%| | 189/189 [01:05<00:00, 2.87it/s, loss: 1.686]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.432876607728383

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.49it/s, acc: 0.489 f1:
0.489]

Epoch accuracy on dev: 0.489
Epoch f1 on dev: 0.489
Epoch: 11 of 100

train batches: 100%| | 189/189 [01:05<00:00, 2.87it/s, loss: 3.147]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.4248313714587497
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.57it/s, acc: 0.492 f1:
0.492]

Epoch accuracy on dev: 0.491
Epoch f1 on dev: 0.491
Epoch: 12 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.84it/s, loss: 1.740]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.42172167130879
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.41it/s, acc: 0.493 f1:
0.493]

Epoch accuracy on dev: 0.492
Epoch f1 on dev: 0.492
Epoch: 13 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.85it/s, loss: 3.059]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.4060298715318953
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.75it/s, acc: 0.493 f1:
0.493]

Epoch accuracy on dev: 0.493
Epoch f1 on dev: 0.493
Epoch: 14 of 100

```
train batches: 100%|      | 189/189 [01:06<00:00,  2.86it/s, loss: 1.688]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.378011323156811
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.43it/s,  acc: 0.493 f1:
0.493]

Epoch accuracy on dev: 0.493
Epoch f1 on dev: 0.493
Epoch: 15 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.86it/s, loss: 3.096]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.368182281653086
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.43it/s,  acc: 0.495 f1:
0.495]

Epoch accuracy on dev: 0.494
Epoch f1 on dev: 0.494
Epoch: 16 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.83it/s, loss: 3.162]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.375419453022972

dev batches: 100%|      | 23/23 [00:02<00:00, 10.50it/s,  acc: 0.494 f1:
0.494]

Epoch accuracy on dev: 0.495
Epoch f1 on dev: 0.495
Epoch: 17 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.85it/s, loss: 3.138]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.3562527015095664
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.46it/s,  acc: 0.495 f1:
0.495]

Epoch accuracy on dev: 0.495
Epoch f1 on dev: 0.495
```

Epoch: 18 of 100

train batches: 100%| | 189/189 [01:05<00:00, 2.89it/s, loss: 1.771]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.327717604145171
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 10.25it/s, acc: 0.496 f1:
0.496]

Epoch accuracy on dev: 0.496
Epoch f1 on dev: 0.496
Epoch: 19 of 100

train batches: 100%| | 189/189 [01:07<00:00, 2.81it/s, loss: 2.814]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.3161519444178023
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.72it/s, acc: 0.498 f1:
0.498]

Epoch accuracy on dev: 0.497
Epoch f1 on dev: 0.497
Epoch: 20 of 100

train batches: 100%| | 189/189 [01:07<00:00, 2.80it/s, loss: 1.746]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.2987569561080328
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.57it/s, acc: 0.499 f1:
0.499]

Epoch accuracy on dev: 0.499
Epoch f1 on dev: 0.499
Epoch: 21 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.86it/s, loss: 3.013]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.285875787810674
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00,  9.66it/s,  acc: 0.500 f1:
0.500]
```

Epoch accuracy on dev: 0.499
Epoch f1 on dev: 0.499
Epoch: 22 of 100

```
train batches: 100%|     | 189/189 [01:06<00:00,  2.84it/s, loss: 3.119]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.263201404185522
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00,  9.87it/s,  acc: 0.500 f1:
0.500]
```

Epoch accuracy on dev: 0.500
Epoch f1 on dev: 0.500
Epoch: 23 of 100

```
train batches: 100%|     | 189/189 [01:06<00:00,  2.82it/s, loss: 1.900]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.2322375812227766
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00, 10.00it/s,  acc: 0.501 f1:
0.501]
```

Epoch accuracy on dev: 0.501
Epoch f1 on dev: 0.501
Epoch: 24 of 100

```
train batches: 100%|     | 189/189 [01:07<00:00,  2.80it/s, loss: 3.208]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.222778018504854
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00,  9.50it/s,  acc: 0.502 f1:
0.502]
```

Epoch accuracy on dev: 0.501
Epoch f1 on dev: 0.501
Epoch: 25 of 100

```
train batches: 100%|     | 189/189 [01:07<00:00,  2.81it/s, loss: 1.672]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.195541304255289
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.10it/s, acc: 0.503 f1:
0.503]

Epoch accuracy on dev: 0.502
Epoch f1 on dev: 0.502
Epoch: 26 of 100

train batches: 100%| | 189/189 [01:08<00:00, 2.76it/s, loss: 2.886]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.2124041044522844

dev batches: 100%| | 23/23 [00:02<00:00, 10.01it/s, acc: 0.504 f1:
0.504]

Epoch accuracy on dev: 0.503
Epoch f1 on dev: 0.503
Epoch: 27 of 100

train batches: 100%| | 189/189 [01:07<00:00, 2.81it/s, loss: 1.894]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.1702942517068653
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.29it/s, acc: 0.502 f1:
0.502]

Epoch accuracy on dev: 0.503
Epoch f1 on dev: 0.503
Epoch: 28 of 100

train batches: 100%| | 189/189 [01:07<00:00, 2.82it/s, loss: 1.671]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.1627806084496632
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.50it/s, acc: 0.503 f1:
0.503]

Epoch accuracy on dev: 0.503
Epoch f1 on dev: 0.503
Epoch: 29 of 100

```
train batches: 100%|        | 189/189 [01:07<00:00,  2.81it/s, loss: 2.319]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.147326547948141
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|        | 23/23 [00:02<00:00,  9.75it/s,  acc: 0.503 f1:
0.503]
```

Epoch accuracy on dev: 0.503
Epoch f1 on dev: 0.503
Epoch: 30 of 100


```
train batches: 100%|        | 189/189 [01:06<00:00,  2.82it/s, loss: 3.147]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.137194598477984
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|        | 23/23 [00:02<00:00,  9.40it/s,  acc: 0.502 f1:
0.502]
```

Epoch accuracy on dev: 0.502
Epoch f1 on dev: 0.502
Epoch: 31 of 100


```
train batches: 100%|        | 189/189 [01:06<00:00,  2.84it/s, loss: 4.571]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.1235754177683877
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|        | 23/23 [00:02<00:00,  9.71it/s,  acc: 0.503 f1:
0.503]
```

Epoch accuracy on dev: 0.503
Epoch f1 on dev: 0.503
Epoch: 32 of 100


```
train batches: 100%|        | 189/189 [01:06<00:00,  2.86it/s, loss: 1.730]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.096411606622121
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|        | 23/23 [00:02<00:00,  9.15it/s,  acc: 0.503 f1:
0.503]
```

Epoch accuracy on dev: 0.503
Epoch f1 on dev: 0.503
Epoch: 33 of 100

train batches: 100%| | 189/189 [01:05<00:00, 2.87it/s, loss: 2.017]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0882263334970625
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.38it/s, acc: 0.504 f1:
0.504]

Epoch accuracy on dev: 0.504
Epoch f1 on dev: 0.504
Epoch: 34 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.82it/s, loss: 1.657]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.096682635564653

dev batches: 100%| | 23/23 [00:02<00:00, 9.57it/s, acc: 0.504 f1:
0.504]

Epoch accuracy on dev: 0.504
Epoch f1 on dev: 0.504
Epoch: 35 of 100

train batches: 100%| | 189/189 [01:07<00:00, 2.82it/s, loss: 1.655]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.079203950980353
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.52it/s, acc: 0.504 f1:
0.504]

Epoch accuracy on dev: 0.504
Epoch f1 on dev: 0.504
Epoch: 36 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.85it/s, loss: 1.656]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0657353949925255
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00, 10.09it/s,  acc: 0.504 f1:
0.504]

Epoch accuracy on dev: 0.504
Epoch f1 on dev: 0.504
Epoch: 37 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.84it/s, loss: 3.150]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0810750770190407

dev batches: 100%|      | 23/23 [00:02<00:00,  9.51it/s,  acc: 0.504 f1:
0.504]

Epoch accuracy on dev: 0.504
Epoch f1 on dev: 0.504
Epoch: 38 of 100


train batches: 100%|      | 189/189 [01:07<00:00,  2.79it/s, loss: 1.727]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.062461595686655
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.39it/s,  acc: 0.505 f1:
0.505]

Epoch accuracy on dev: 0.505
Epoch f1 on dev: 0.505
Epoch: 39 of 100


train batches: 100%|      | 189/189 [01:07<00:00,  2.81it/s, loss: 3.069]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0526977436883107
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.30it/s,  acc: 0.504 f1:
0.504]

Epoch accuracy on dev: 0.505
Epoch f1 on dev: 0.505
Epoch: 40 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.84it/s, loss: 1.661]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0544499546762496
```

```
dev batches: 100%|      | 23/23 [00:02<00:00, 10.83it/s,  acc: 0.504 f1:
0.504]
```

Epoch accuracy on dev: 0.504
Epoch f1 on dev: 0.504
Epoch: 41 of 100


```
train batches: 100%|      | 189/189 [01:06<00:00,  2.83it/s, loss: 1.669]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.031692077243139
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00,  9.43it/s,  acc: 0.505 f1:
0.505]
```

Epoch accuracy on dev: 0.505
Epoch f1 on dev: 0.505
Epoch: 42 of 100


```
train batches: 100%|      | 189/189 [01:07<00:00,  2.81it/s, loss: 1.665]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.015584131081899
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00,  8.80it/s,  acc: 0.505 f1:
0.505]
```

Epoch accuracy on dev: 0.505
Epoch f1 on dev: 0.505
Epoch: 43 of 100


```
train batches: 100%|      | 189/189 [01:08<00:00,  2.77it/s, loss: 3.737]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.036620264961606

```
dev batches: 100%|      | 23/23 [00:02<00:00, 10.59it/s,  acc: 0.506 f1:
0.506]
```

Epoch accuracy on dev: 0.506
Epoch f1 on dev: 0.506
Epoch: 44 of 100


```
train batches: 100%|      | 189/189 [01:08<00:00,  2.78it/s, loss: 1.706]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0634647711874945

```
dev batches: 100%|     | 23/23 [00:02<00:00, 10.25it/s,  acc: 0.506 f1:
0.506]
```

Epoch accuracy on dev: 0.506
Epoch f1 on dev: 0.506
Epoch: 45 of 100

```
train batches: 100%|     | 189/189 [01:07<00:00,  2.79it/s, loss: 3.054]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0068282059260776
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|     | 23/23 [00:02<00:00,  8.96it/s,  acc: 0.507 f1:
0.507]
```

Epoch accuracy on dev: 0.507
Epoch f1 on dev: 0.507
Epoch: 46 of 100

```
train batches: 100%|     | 189/189 [01:07<00:00,  2.80it/s, loss: 3.154]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.01315936115053

```
dev batches: 100%|     | 23/23 [00:02<00:00, 10.80it/s,  acc: 0.507 f1:
0.507]
```

Epoch accuracy on dev: 0.507
Epoch f1 on dev: 0.507
Epoch: 47 of 100

```
train batches: 100%|     | 189/189 [01:07<00:00,  2.82it/s, loss: 1.655]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9994330103435214
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|     | 23/23 [00:02<00:00,  9.44it/s,  acc: 0.508 f1:
0.508]
```

Epoch accuracy on dev: 0.508
Epoch f1 on dev: 0.508
Epoch: 48 of 100

```
train batches: 100%|     | 189/189 [01:06<00:00,  2.84it/s, loss: 1.657]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0010504419841464

```
dev batches: 100%|      | 23/23 [00:02<00:00, 10.57it/s,  acc: 0.509 f1:
0.509]

Epoch accuracy on dev: 0.508
Epoch f1 on dev: 0.508
Epoch: 49 of 100


train batches: 100%|      | 189/189 [01:07<00:00,  2.81it/s, loss: 1.656]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.012135068575541

dev batches: 100%|      | 23/23 [00:02<00:00,  9.26it/s,  acc: 0.509 f1:
0.509]

Epoch accuracy on dev: 0.509
Epoch f1 on dev: 0.509
Epoch: 50 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.86it/s, loss: 3.123]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0059115319024947
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.17it/s,  acc: 0.509 f1:
0.509]

Epoch accuracy on dev: 0.509
Epoch f1 on dev: 0.509
Epoch: 51 of 100


train batches: 100%|      | 189/189 [01:05<00:00,  2.86it/s, loss: 3.023]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 2.0086637262314087

dev batches: 100%|      | 23/23 [00:02<00:00, 10.86it/s,  acc: 0.509 f1:
0.509]

Epoch accuracy on dev: 0.509
Epoch f1 on dev: 0.509
Epoch: 52 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.84it/s, loss: 3.026]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9916304094450814
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth
```

```
dev batches: 100%|     | 23/23 [00:02<00:00,  9.80it/s,  acc: 0.509 f1:
0.509]

Epoch accuracy on dev: 0.509
Epoch f1 on dev: 0.509
Epoch: 53 of 100


train batches: 100%|     | 189/189 [01:06<00:00,  2.86it/s, loss: 1.660]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9997245839663915

dev batches: 100%|     | 23/23 [00:02<00:00, 10.68it/s,  acc: 0.510 f1:
0.510]

Epoch accuracy on dev: 0.510
Epoch f1 on dev: 0.510
Epoch: 54 of 100


train batches: 100%|     | 189/189 [01:05<00:00,  2.90it/s, loss: 1.659]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9856459488944402
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|     | 23/23 [00:02<00:00,  9.42it/s,  acc: 0.510 f1:
0.510]

Epoch accuracy on dev: 0.510
Epoch f1 on dev: 0.510
Epoch: 55 of 100


train batches: 100%|     | 189/189 [01:07<00:00,  2.81it/s, loss: 1.679]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.984319477800339
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|     | 23/23 [00:02<00:00,  9.51it/s,  acc: 0.510 f1:
0.510]

Epoch accuracy on dev: 0.510
Epoch f1 on dev: 0.510
Epoch: 56 of 100


train batches: 100%|     | 189/189 [01:05<00:00,  2.87it/s, loss: 1.657]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9826118350028992
```

Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 8.98it/s, acc: 0.511 f1:
0.511]

Epoch accuracy on dev: 0.510
Epoch f1 on dev: 0.510
Epoch: 57 of 100

train batches: 100%| | 189/189 [01:05<00:00, 2.87it/s, loss: 3.135]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9910619873849174

dev batches: 100%| | 23/23 [00:02<00:00, 10.54it/s, acc: 0.511 f1:
0.511]

Epoch accuracy on dev: 0.511
Epoch f1 on dev: 0.511
Epoch: 58 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.84it/s, loss: 3.152]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9848860208950345
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.40it/s, acc: 0.512 f1:
0.512]

Epoch accuracy on dev: 0.511
Epoch f1 on dev: 0.511
Epoch: 59 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.82it/s, loss: 1.690]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9765426213779147
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.50it/s, acc: 0.512 f1:
0.512]

Epoch accuracy on dev: 0.512
Epoch f1 on dev: 0.512
Epoch: 60 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.84it/s, loss: 1.680]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9795710898581005

dev batches: 100%| | 23/23 [00:02<00:00, 11.00it/s, acc: 0.512 f1:
0.512]

Epoch accuracy on dev: 0.512
Epoch f1 on dev: 0.512
Epoch: 61 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.83it/s, loss: 2.487]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9782548679245844
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 10.62it/s, acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.512
Epoch f1 on dev: 0.512
Epoch: 62 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.84it/s, loss: 1.657]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9855035448831224

dev batches: 100%| | 23/23 [00:02<00:00, 9.52it/s, acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 63 of 100

train batches: 100%| | 189/189 [01:04<00:00, 2.93it/s, loss: 1.681]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9669001953942435
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.28it/s, acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 64 of 100

train batches: 100%| | 189/189 [01:04<00:00, 2.92it/s, loss: 1.660]

```
Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9753160713210938

dev batches: 100%|       | 23/23 [00:02<00:00,  9.22it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 65 of 100


train batches: 100%|       | 189/189 [01:05<00:00,  2.90it/s, loss: 1.818]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9716898212357172
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|       | 23/23 [00:02<00:00,  9.59it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 66 of 100


train batches: 100%|       | 189/189 [01:04<00:00,  2.92it/s, loss: 1.663]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9686179955800374
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|       | 23/23 [00:02<00:00,  9.68it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 67 of 100


train batches: 100%|       | 189/189 [01:04<00:00,  2.93it/s, loss: 1.657]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9615404331494892
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|       | 23/23 [00:02<00:00,  9.84it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 68 of 100
```

```
train batches: 100%|      | 189/189 [01:05<00:00,  2.88it/s, loss: 3.153]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9839915207454137

dev batches: 100%|      | 23/23 [00:02<00:00, 11.10it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 69 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.84it/s, loss: 1.684]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9638444412322271
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.24it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 70 of 100


train batches: 100%|      | 189/189 [01:05<00:00,  2.86it/s, loss: 2.431]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9654856267429532

dev batches: 100%|      | 23/23 [00:02<00:00, 10.45it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 71 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.85it/s, loss: 3.152]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9757023794310433

dev batches: 100%|      | 23/23 [00:02<00:00, 11.02it/s,  acc: 0.512 f1:
0.512]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 72 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.84it/s, loss: 1.664]
```

```
Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9663139070783342
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.54it/s,  acc: 0.512 f1:
0.512]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 73 of 100


train batches: 100%|      | 189/189 [01:05<00:00,  2.87it/s, loss: 1.669]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9525437638873147
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00, 10.25it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 74 of 100


train batches: 100%|      | 189/189 [01:07<00:00,  2.81it/s, loss: 1.655]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9593601633631994

dev batches: 100%|      | 23/23 [00:02<00:00, 10.04it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 75 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.86it/s, loss: 1.655]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9567487296603976
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00,  9.53it/s,  acc: 0.513 f1:
0.513]

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 76 of 100
```

```
train batches: 100%|      | 189/189 [01:06<00:00,  2.85it/s, loss: 1.655]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9516242685772123
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00,  9.89it/s,  acc: 0.513 f1:
0.513]
```

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 77 of 100


```
train batches: 100%|      | 189/189 [01:06<00:00,  2.86it/s, loss: 1.656]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.960463855947767

```
dev batches: 100%|      | 23/23 [00:02<00:00, 10.71it/s,  acc: 0.514 f1:
0.514]
```

Epoch accuracy on dev: 0.513
Epoch f1 on dev: 0.513
Epoch: 78 of 100


```
train batches: 100%|      | 189/189 [01:06<00:00,  2.85it/s, loss: 3.154]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9589133735687014
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00,  9.66it/s,  acc: 0.513 f1:
0.513]
```

Epoch accuracy on dev: 0.514
Epoch f1 on dev: 0.514
Epoch: 79 of 100


```
train batches: 100%|      | 189/189 [01:06<00:00,  2.86it/s, loss: 1.683]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9501091743272447
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|      | 23/23 [00:02<00:00,  9.49it/s,  acc: 0.514 f1:
0.514]
```

Epoch accuracy on dev: 0.514
Epoch f1 on dev: 0.514

```

Epoch: 80 of 100

train batches: 100%|     | 189/189 [01:07<00:00,  2.81it/s, loss: 3.094]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9630927292127458

dev batches: 100%|     | 23/23 [00:02<00:00, 11.14it/s,  acc: 0.514 f1:
0.514]

Epoch accuracy on dev: 0.514
Epoch f1 on dev: 0.514
Epoch: 81 of 100

train batches: 100%|     | 189/189 [01:06<00:00,  2.82it/s, loss: 1.658]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9448664443833488
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|     | 23/23 [00:02<00:00,  9.22it/s,  acc: 0.514 f1:
0.514]

Epoch accuracy on dev: 0.514
Epoch f1 on dev: 0.514
Epoch: 82 of 100

train batches: 100%|     | 189/189 [01:06<00:00,  2.85it/s, loss: 1.655]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.953756095871093

dev batches: 100%|     | 23/23 [00:02<00:00, 11.11it/s,  acc: 0.514 f1:
0.514]

Epoch accuracy on dev: 0.514
Epoch f1 on dev: 0.514
Epoch: 83 of 100

train batches: 100%|     | 189/189 [01:07<00:00,  2.81it/s, loss: 1.663]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9523620274331834
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|     | 23/23 [00:02<00:00,  9.05it/s,  acc: 0.515 f1:
0.515]

Epoch accuracy on dev: 0.514
Epoch f1 on dev: 0.514

```
Epoch: 84 of 100

train batches: 100%| | 189/189 [01:07<00:00, 2.79it/s, loss: 1.655]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9495165376436143
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.26it/s, acc: 0.515 f1:
0.515]

Epoch accuracy on dev: 0.515
Epoch f1 on dev: 0.515
Epoch: 85 of 100

train batches: 100%| | 189/189 [01:05<00:00, 2.88it/s, loss: 1.657]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9528056865646726

dev batches: 100%| | 23/23 [00:02<00:00, 10.79it/s, acc: 0.515 f1:
0.515]

Epoch accuracy on dev: 0.515
Epoch f1 on dev: 0.515
Epoch: 86 of 100

train batches: 100%| | 189/189 [01:05<00:00, 2.88it/s, loss: 1.663]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.949629462900616
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.66it/s, acc: 0.515 f1:
0.515]

Epoch accuracy on dev: 0.515
Epoch f1 on dev: 0.515
Epoch: 87 of 100

train batches: 100%| | 189/189 [01:06<00:00, 2.83it/s, loss: 1.657]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9494493755083235
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%| | 23/23 [00:02<00:00, 9.70it/s, acc: 0.516 f1:
0.516]
```

Epoch accuracy on dev: 0.515
Epoch f1 on dev: 0.515
Epoch: 88 of 100


train batches: 100%|    | 189/189 [01:06<00:00, 2.85it/s, loss: 1.655]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9538799354008265

dev batches: 100%|    | 23/23 [00:02<00:00, 10.60it/s, acc: 0.516 f1:
0.516]

Epoch accuracy on dev: 0.516
Epoch f1 on dev: 0.516
Epoch: 89 of 100


train batches: 100%|    | 189/189 [01:07<00:00, 2.79it/s, loss: 1.655]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.950244543098268
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|    | 23/23 [00:02<00:00, 9.36it/s, acc: 0.516 f1:
0.516]

Epoch accuracy on dev: 0.516
Epoch f1 on dev: 0.516
Epoch: 90 of 100


train batches: 100%|    | 189/189 [01:06<00:00, 2.86it/s, loss: 3.013]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9495056545923626
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|    | 23/23 [00:02<00:00, 9.07it/s, acc: 0.516 f1:
0.516]

Epoch accuracy on dev: 0.516
Epoch f1 on dev: 0.516
Epoch: 91 of 100


train batches: 100%|    | 189/189 [01:08<00:00, 2.77it/s, loss: 1.688]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9399433514428517
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|     | 23/23 [00:02<00:00,  9.10it/s,  acc: 0.516 f1:
0.516]
```

Epoch accuracy on dev: 0.516
Epoch f1 on dev: 0.516
Epoch: 92 of 100

```
train batches: 100%|     | 189/189 [01:06<00:00,  2.84it/s, loss: 1.677]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9483399173570057

```
dev batches: 100%|     | 23/23 [00:02<00:00, 10.97it/s,  acc: 0.516 f1:
0.516]
```

Epoch accuracy on dev: 0.516
Epoch f1 on dev: 0.516
Epoch: 93 of 100

```
train batches: 100%|     | 189/189 [01:08<00:00,  2.78it/s, loss: 1.660]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.950664408623226

```
dev batches: 100%|     | 23/23 [00:02<00:00,  9.40it/s,  acc: 0.517 f1:
0.517]
```

Epoch accuracy on dev: 0.516
Epoch f1 on dev: 0.516
Epoch: 94 of 100

```
train batches: 100%|     | 189/189 [01:06<00:00,  2.84it/s, loss: 1.822]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9405182070202298
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

```
dev batches: 100%|     | 23/23 [00:02<00:00,  9.26it/s,  acc: 0.516 f1:
0.516]
```

Epoch accuracy on dev: 0.517
Epoch f1 on dev: 0.517
Epoch: 95 of 100

```
train batches: 100%|     | 189/189 [01:06<00:00,  2.86it/s, loss: 1.660]
```

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9440210499460735

```
dev batches: 100%|     | 23/23 [00:02<00:00, 10.65it/s,  acc: 0.516 f1:
0.516]
```

```
Epoch accuracy on dev: 0.516
Epoch f1 on dev: 0.516
Epoch: 96 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.83it/s, loss: 3.154]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9484724970090956

dev batches: 100%|      | 23/23 [00:02<00:00, 10.89it/s,  acc: 0.517 f1:
0.517]

Epoch accuracy on dev: 0.517
Epoch f1 on dev: 0.517
Epoch: 97 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.85it/s, loss: 1.655]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9443574926209828
Saved model to: /Users/johnsonzhou/git/comp90042-project/result/models/model_04_
roberta_base_3_class.pth

dev batches: 100%|      | 23/23 [00:02<00:00, 10.23it/s,  acc: 0.517 f1:
0.517]

Epoch accuracy on dev: 0.517
Epoch f1 on dev: 0.517
Epoch: 98 of 100


train batches: 100%|      | 189/189 [01:06<00:00,  2.84it/s, loss: 1.673]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9446840589008634

dev batches: 100%|      | 23/23 [00:02<00:00, 10.65it/s,  acc: 0.517 f1:
0.517]

Epoch accuracy on dev: 0.517
Epoch f1 on dev: 0.517
Epoch: 99 of 100


train batches: 100%|      | 189/189 [01:07<00:00,  2.82it/s, loss: 1.655]

Adjusting learning rate of group 0 to 2.0000e-06.
Average epoch loss: 1.9458723172308907

dev batches: 100%|      | 23/23 [00:02<00:00,  9.81it/s,  acc: 0.517 f1:
0.517]
```

```
Epoch accuracy on dev: 0.517
Epoch f1 on dev: 0.517
Done!
```