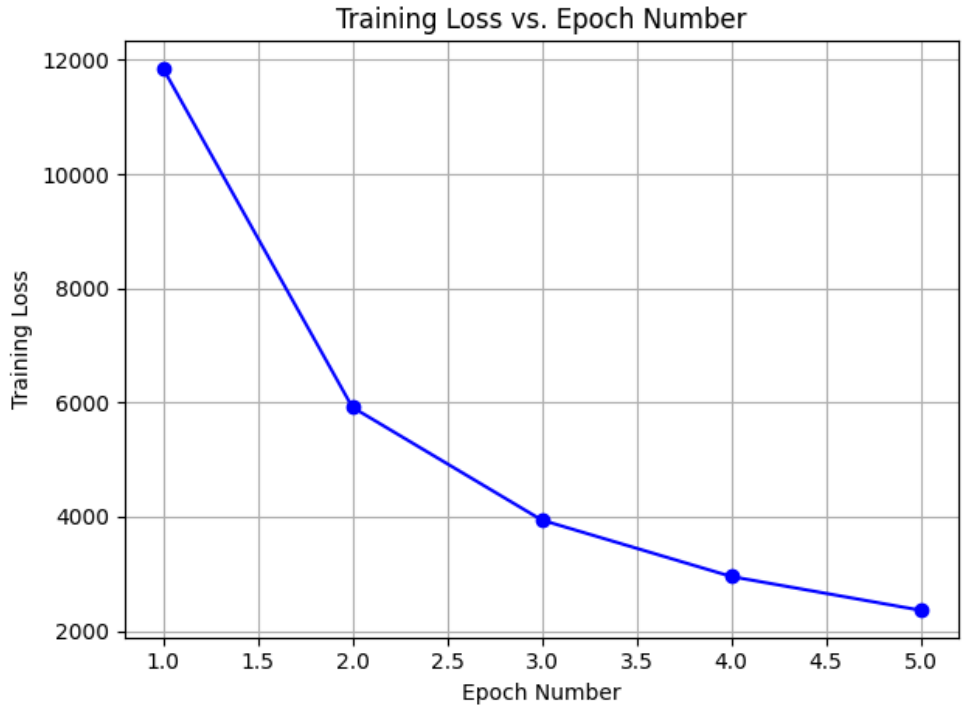
NLP ASSIGNMENT 4 REPORT

TASK-1: Emotion Recognition in Conversations Task (ERC):

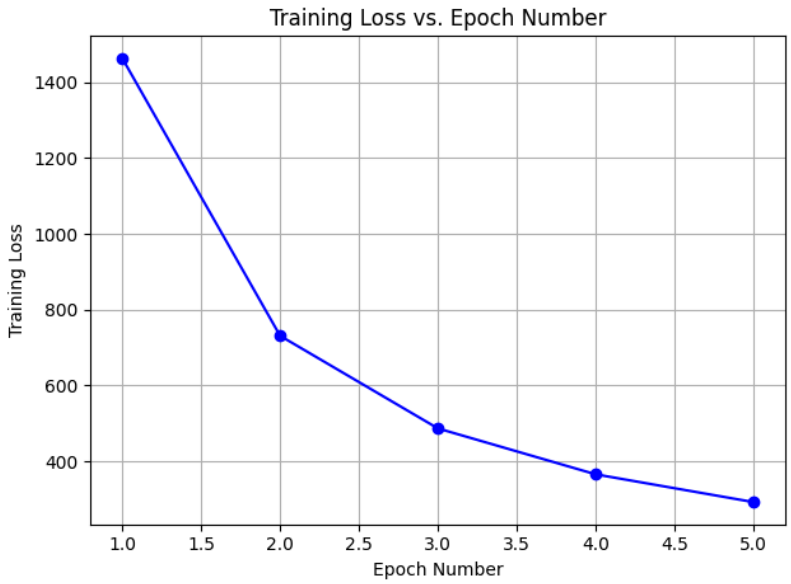
M1: Finetuning the pretrained BERT ('bert-base-uncased'):

In this setup, we decided to finetune the pretrained BERT model ('bert-base-uncased') available on Huggingface. We converted the given data into lists of various attributes such as speakers, utterances, emotions and so on. For this task, we have ignored the speakers' list as far as the model training is concerned, but we required it to determine if there has been an emotion flip or not. We do NOT need a model to detect whether an emotion flip has occurred or not, and this has been determined via an algorithm. The model here only takes in a list of utterances, and their corresponding emotions during training, and given the utterances of the test dataset, it predicts the emotions for those utterances.
As far as training is concerned, here we have applied BERT embedding, and passed the lists through a DataLoader. The macro F1 score has been predicted on the validation dataset as the test dataset was not given.
The Google Drive LInk to the trained models:
https://drive.google.com/drive/folders/1hHkhY_XzyL9y2kCWopk9l4K8lXIZeLH5?usp=sharing

Training Loss vs Epoch Curve:



Validation Loss vs Epoch Curve:



Macro-F1 Score reported: 0.5327

```
from sklearn.metrics import f1_score
macro_f1 = f1_score(finalLabels, finalPredictions, average='macro')
print(f'Macro F1 Score: {macro_f1:.4f}')
```

```
Macro F1 Score: 0.5327
```

We have used BERT for this task because it is pretrained on NSP tasks, and understands the relatedness between sentences. Utterances are nothing but sentences, and their relatedness is important for understanding and predicting emotions. Thus, we have fine tuned the pretrained BERT model for this task.

For M2:
We have finetuned GPT2 on pretrained BERT embeddings.
[0.012146675317112882, 0.007030130475448328, 0.0055243211553830576, 0.005667378763189467, 0.004112504017630371]

https://drive.google.com/drive/folders/1hI-5e2GhtvmCfxcYGIpJJ-3AIXEw8JZI?usp=drive_link

TASK-2: Emotion Flip Reasoning (EFR):

M3:

In this model we made pairs of target utterances and possible utterances and fed them to pretrained BERT. Then we concatenated the output of BERT with one-hot encoded emotions of the two sentences that were

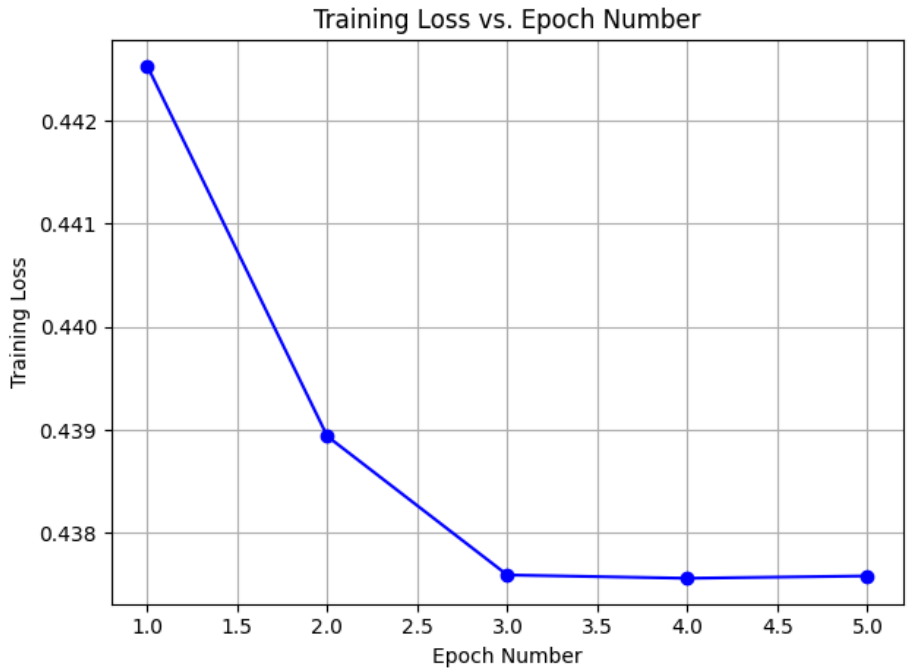
fed to BERT, these one-hot encodings of two emotions were also concatenated with one another. Then this whole tensor is fed to a MLP which has 3 layers(782, 400, 100 neurons respectively). At the end of this MLP is a sigmoid function that will help to predict whether possible utterance is responsible for the flip in the target utterance or not. Also the weights of BERT are updated during training i.e its being fine-tuned for our task.

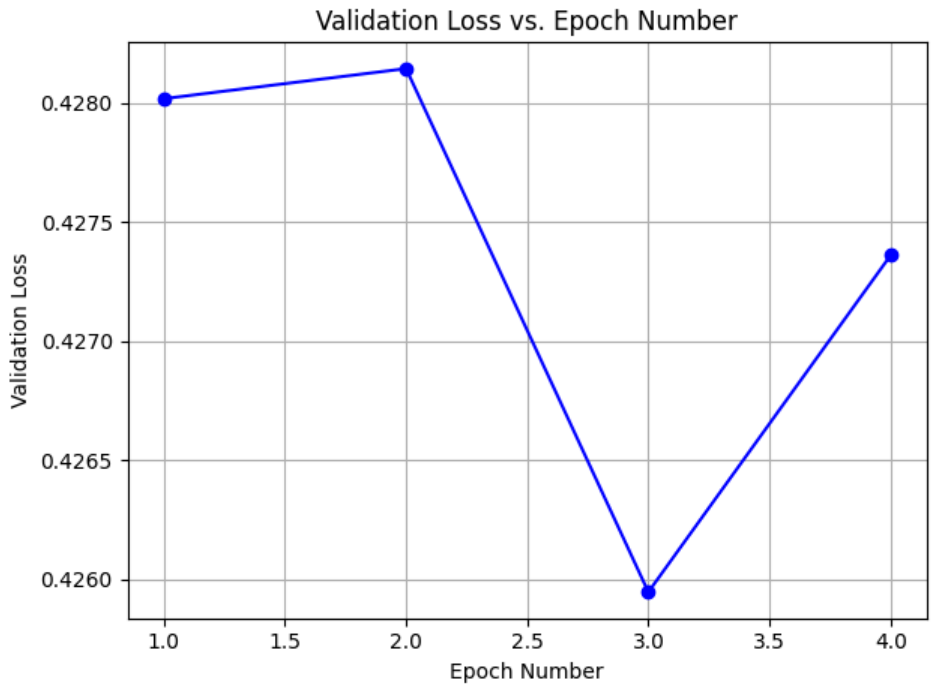
This model is better than then other because it uses pretrained BERT to generate encoding for pair of sentences. Because of pretraining BERT understands english language and semantic meaning of the sentence that we give to it, hence it provides better encodings.

F-1 Score: 0.4589117363224705

Model Link:

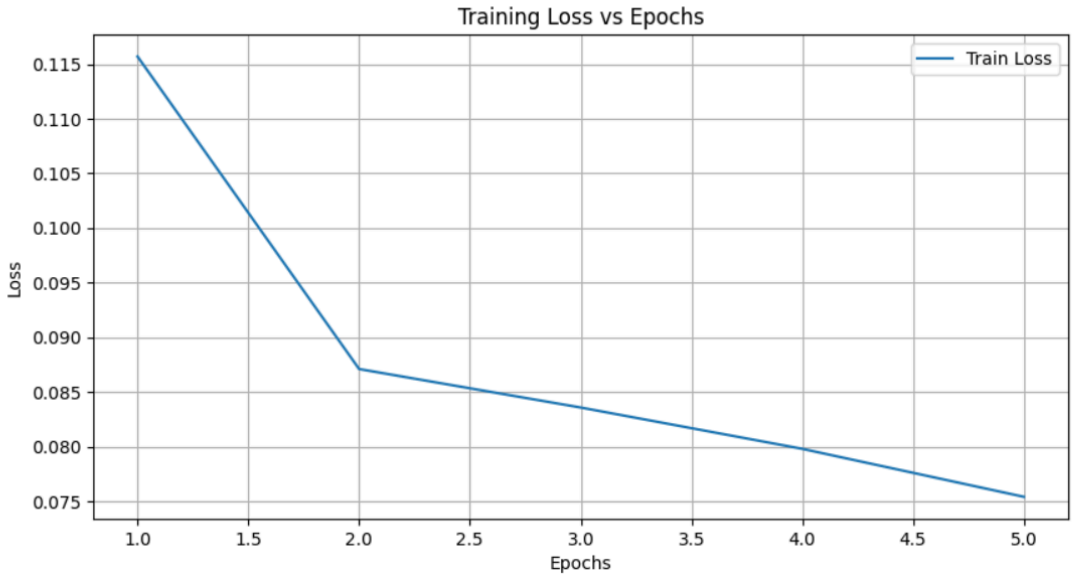
https://drive.google.com/drive/folders/1RV38IqZo_jc8KWZyBUoVr2yFJIQXqw5F?usp=sharing

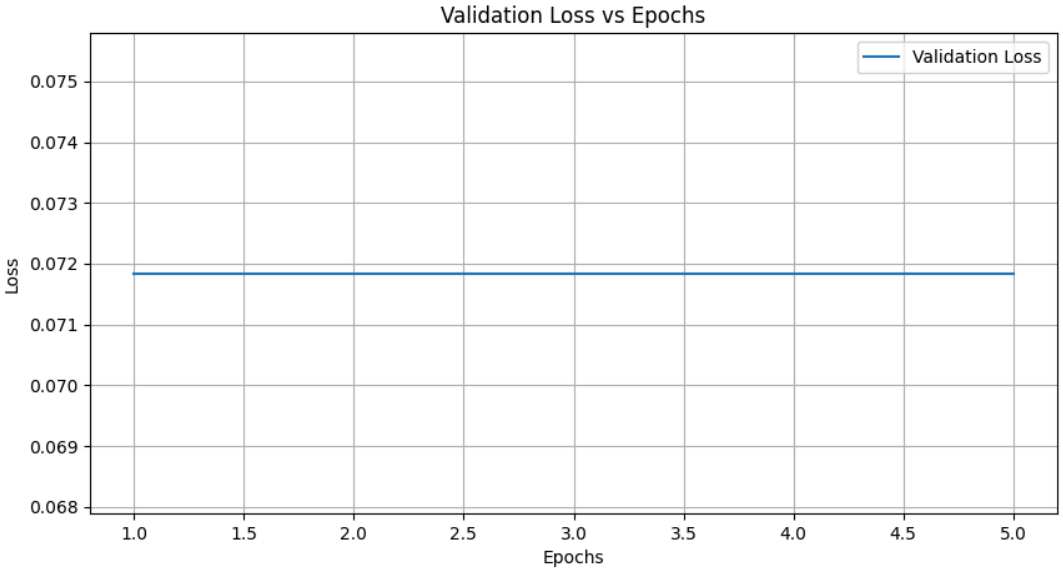




M4

We have used GRU model





Training f1 scores for each epoch
0.02558618289953708
0.029378551461972657
0.03222399342358575
0.03659973823531939
0.04304540243941587
Final f1 score = 0.03336677369196615

validation f1 score
0.054553176788437255
0.054553176788437255
0.054553176788437255
0.054553176788437255
0.054553176788437255
Final f1 score for validation
0.054553176788437255

NLP_A4 - Google Drive
https://drive.google.com/drive/folders/1se2nRpRclbmVk4k7yI1NWnBH9BCQvHe-?usp=drive_link

CONTRIBUTION and TEAM MEMBERS:

M1: Rohan Gupta (2020113)

M2: Jay Saraf (2020438)

M3: Vatsal Chaudhary (2020549)

M4: Harsh (2020061)