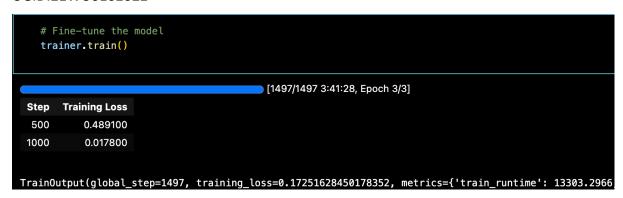Name: Gurram Sai Vinil
UGID:21WU0102022

Social Media Analytics Report

Dataset Used:

To generate summaries for this analysis, we used test.csv in the following dataset:

https://www.kaggle.com/datasets/gowrishankarp/newspaper-text-summarization-cnn-dailymail

- Due to computational constraints, a subset of 500 rows was employed for model training.

Procedure for Obtaining Bert_Score and Rouge_Score:

- Extraction of key components from the dataset was performed. In this case, Articles and Summaries were selected.

- The Flan-T5 model was employed for text summarization.

- Articles were segmented into manageable chunks, and corresponding summaries were generated.

- Bert and Rouge scores were computed by comparing the predicted summaries with the actual summaries.

Fine-Tuning of Language Model (LLM):

For the fine-tuning process, the following dataset was used:

https://www.kaggle.com/datasets/gowrishankarp/newspaper-text-summarization-cnn-dailymail

- A subset of 500 rows was for training, utilizing the T5 Large model.

```
# Fine-tune the model
trainer.train()
```

[1497/1497 3:41:28, Epoch 3/3]

Step	Training Loss
500	0.489100
1000	0.017800

```
TrainOutput(global_step=1497, training_loss=0.17251628450178352, metrics={'train_runtime': 13303.2966
```

Result:

Bert_Score before fine tuning :

```
Final Cumulative BERT Scores:
Precision: 0.7750833338499069
Recall: 0.7775522378683091
F1 Score: 0.7750406149625778
```

Bert_Score After fine tuning :

```
Final Cumulative BERT Scores:
Precision: 0.8082359130970223
Recall: 0.8203413494842086
F1 Score: 0.8137665107876122
```

Rouge_Score before fine tuning:

```
ROUGE Scores:
ROUGE-1: 0.2693911591773547
ROUGE-2: 0.09009539434010885
ROUGE-L: 0.17814120885358134
ROUGE-Lsum: 0.17839670239171068
```

Rouge_Score After fine tuning :

```
ROUGE Scores:
ROUGE-1: 0.38435374702831687
ROUGE-2: 0.17630441663777796
ROUGE-L: 0.27538776593527947
ROUGE-Lsum: 0.33334323339611777
```

Name: Gurram Sai Vinil
UGID:21WU0102022

Bleu_Score before fine tuning:

Average BLEU Score: 0.29171001991417306

Bleu_Score After fine tuning :

Average BLEU Score: 0.29189278199201235