1) Plot accuracy and loss curve
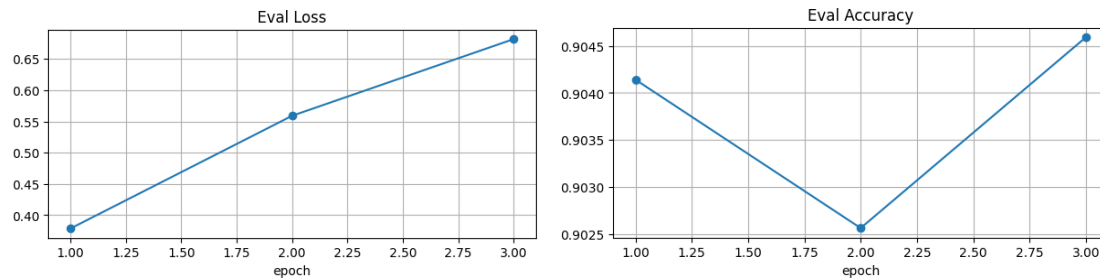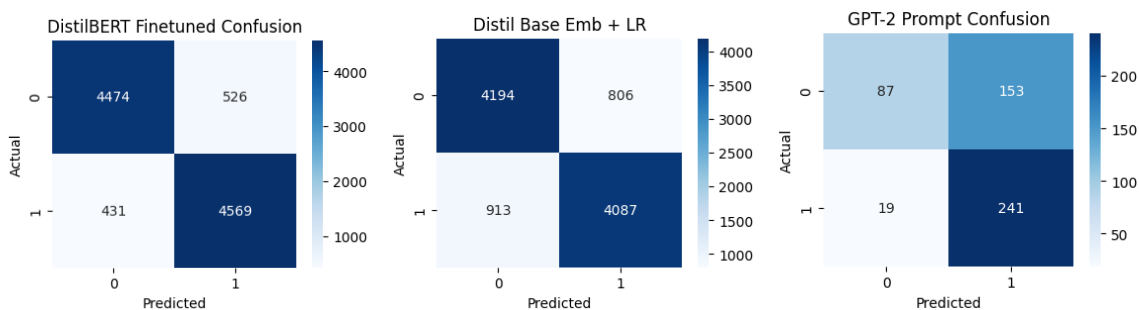

2) Confusion matrix

3) Comparative tables for evaluation matrix

```
Comparative:
                       Model   accuracy   precision     recall         f1
0               TF-IDF + LR     0.9013    0.893354   0.911400   0.902287
1       DistilBERT (finetuned)   0.9043    0.896762   0.913800   0.905201
2   DistilBERT (base emb) + LR   0.8281    0.835275   0.817400   0.826241
3               GPT-2 (prompt)   0.6560    0.611675   0.926923   0.737003
```

4) Report

a. What do the accuracy and loss curves tell you about the fine-tuning process?

Accuracy and loss curves show how well a model learns. For my model, the graph shows an increase of loss over time and a marginal increase in accuracy. This could suggest overfitting or that the learning rate is too high.

b. How does the fine-tuned DistilBERT model compare to the classical ML model? What advantages or limitations do transformers present over classical algorithms?

The DistilBERT model achieves better accuracy than the classical ML model. Transformers can capture word context while classical models only rely on word frequency.

c. What insights can you draw from the confusion matrix? Are there any patterns in the misclassifications?

The confusion matrix for my TF-IDF shows that the majority of the predictions are correctly classified and that there is little bias for false positives and false negatives. The confusion model for GPT-2 shows that there is a bias for false positives as the model predicts that the most of the movies are positive regardless if it is positive or actually negative.

d. Why might the fine-tuned model outperform the base model?

The Fine tuned model is able to adapt to the DistilBERT weights. Because of this, it can learn sentiment.

e. Which model would you recommend for deployment in a real-world scenario, and why?
Consider both performance and efficiency in your answer.

I would use the fine-tuned DistilBERT model because it is the most accurate while also being more efficient than the GTP2 and TF-IDF model.