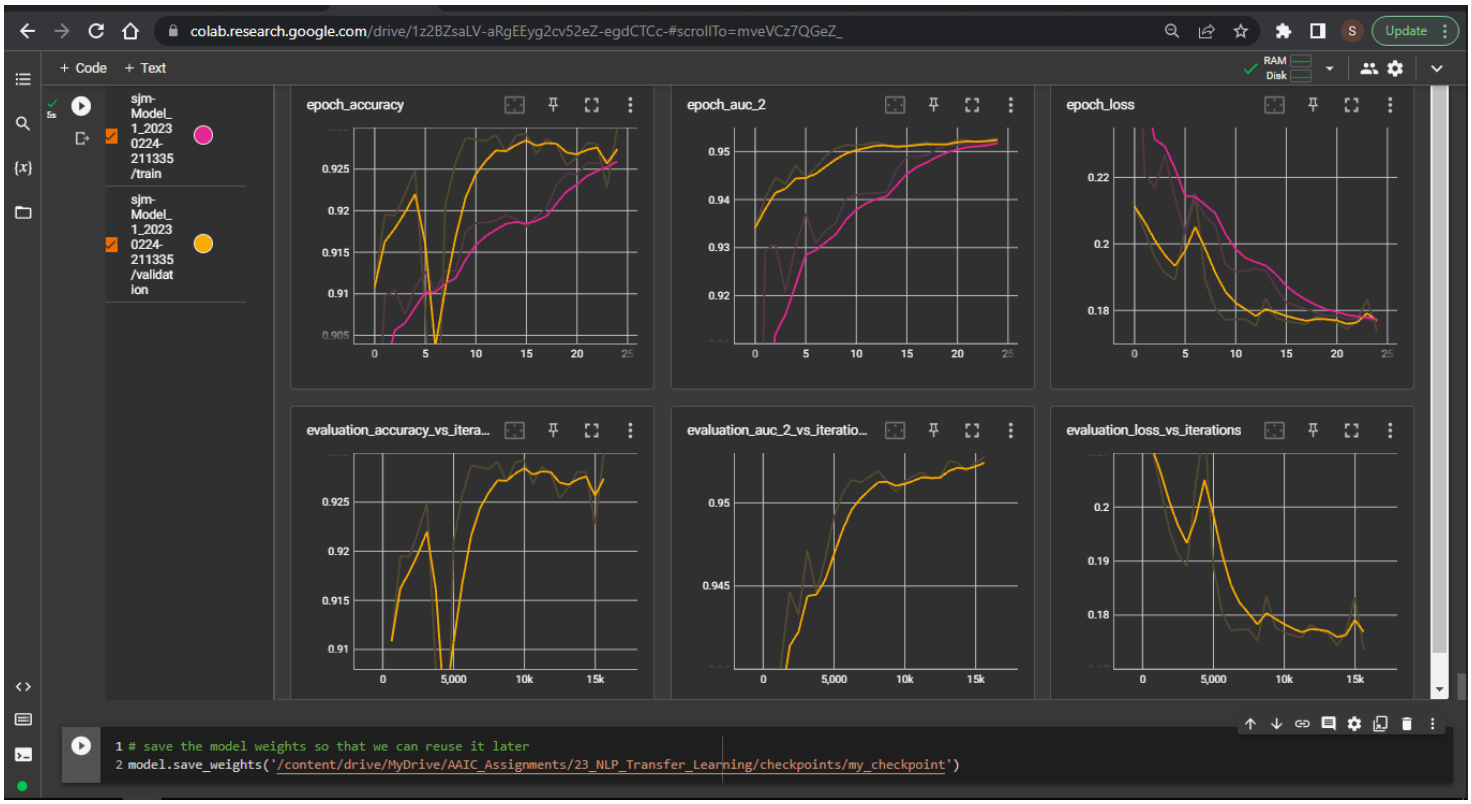
NLP with Transfer Learning

Scalars:



The train accuracy, AUC and loss are improving over the epochs.
The validation/test accuracy first increases then takes a dip and further increases to finally become plateaued.
Same pattern is seen in test AUC which increase over the epochs and gradually plateaus over epochs.

The model's performance is high to begin with and further increases over the epochs. This is due to the fact that BERT produces features that are able to capture the sentiment of the text data.

One thing to notice here is that the validation scores are better than the train scores. One possible reason for this is that the weight initialization that we are starting with is performing better for the validation data rather than the whole of train data. But over time this gap reduces, thereby showing that the model weights are improving with the epochs and the validation scores also increase with it.