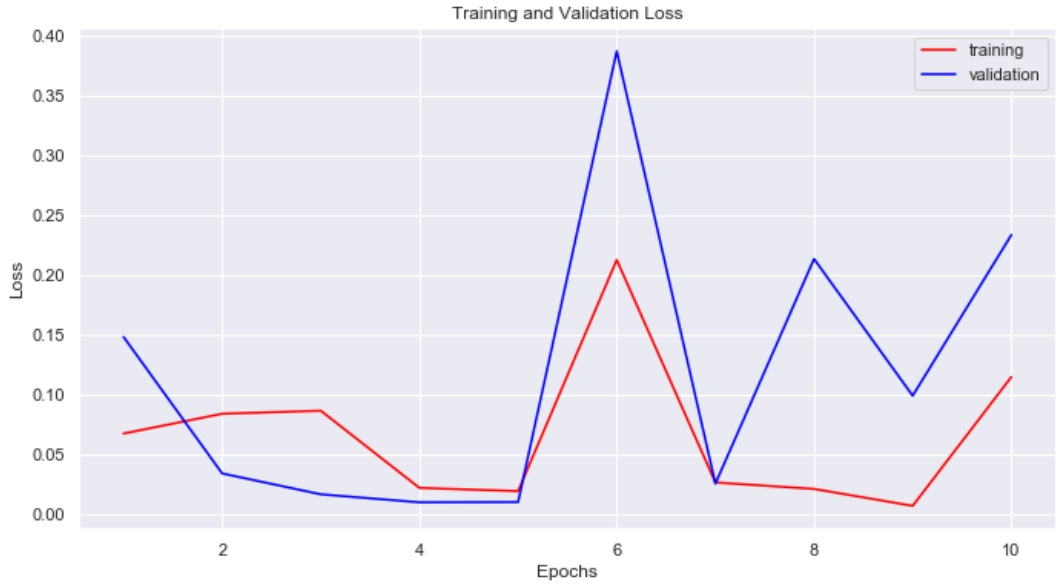
Model Analysis BERT-BASE-UNCASED

Training and Validation BCE-Loss Plot



As we can se at epoch 5 train and validation are low and after that validaton loss is greater as compared to training, So there is overfitting after 5th Epoch

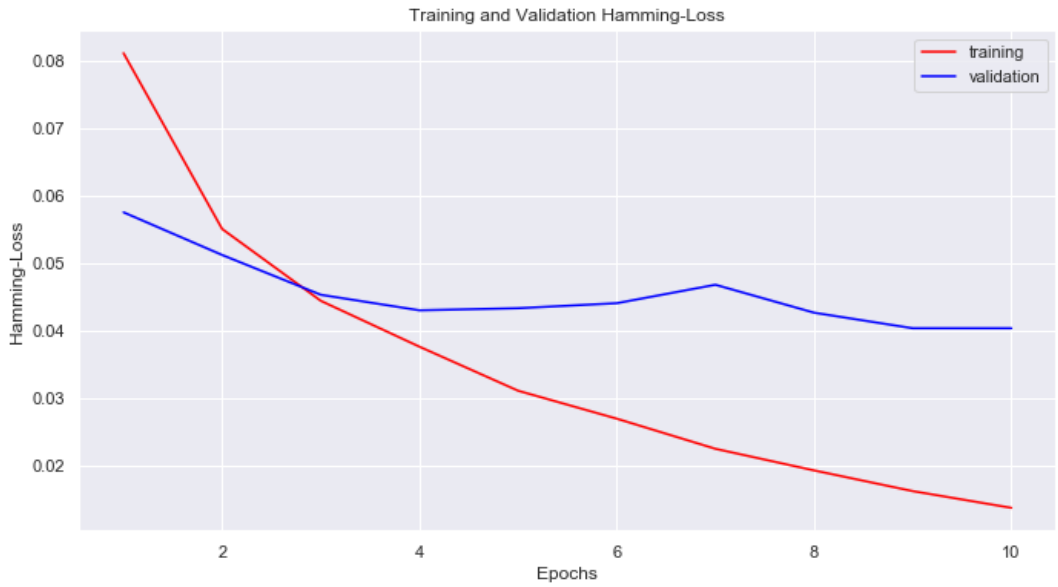
Training and Validation F1-Micro-Score Plot

Here also we can see that after 5th epoch validation F1-Micro-Score is not increasing while training F1-Micro-Score is increasing which shows overfitting after 5th Epoch, Training F1-Micro-Score reached to .90 at 10th Epoch

Training and Validation F1-Macro-Score Plot

Here also we can see that after 5th epoch validation F1-Macro-Score is not increasing while training F1-Macro-Score is increasing which shows overfitting after 5th Epoch, Training F1-Macro-Score reached to .85 at 10th Epoch

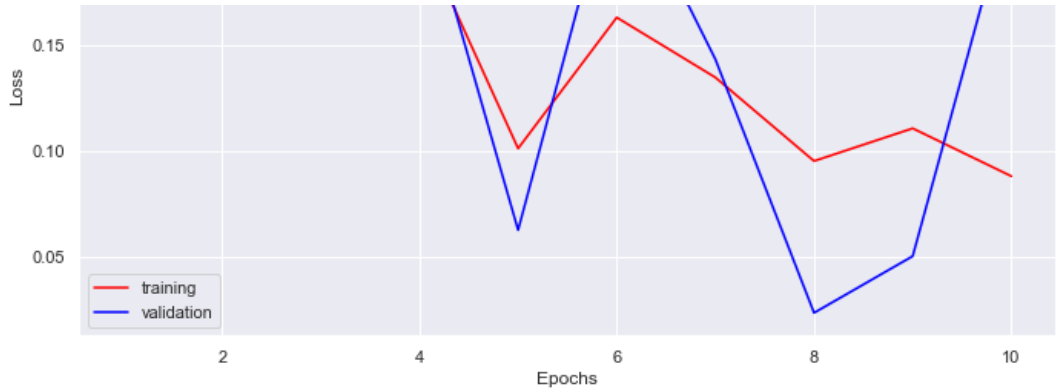
Training and Validation Hamming Loss Plot



As we can se at epoch 5 train and validation are low and after that validaton loss is greater as compared to training, So there is overfitting after 5th Epoch

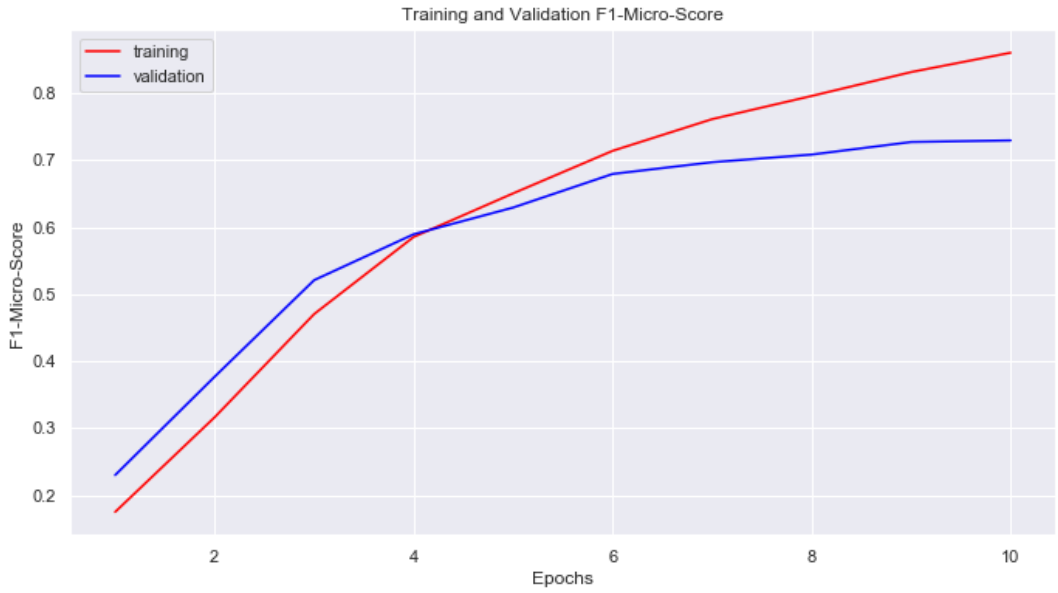
Model Analysis BERT-LARGE-UNCASED

Training and Validation BCE-Loss Plot



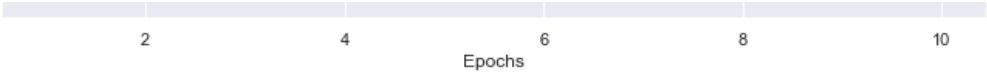
As we can se at epoch 5 train and validation are low and close to each other and after that validaton loss is very low at 8th epoch but training loss is greater so at 5th epoch our model is performing best

Training and Validation F1-Micro-Score Plot



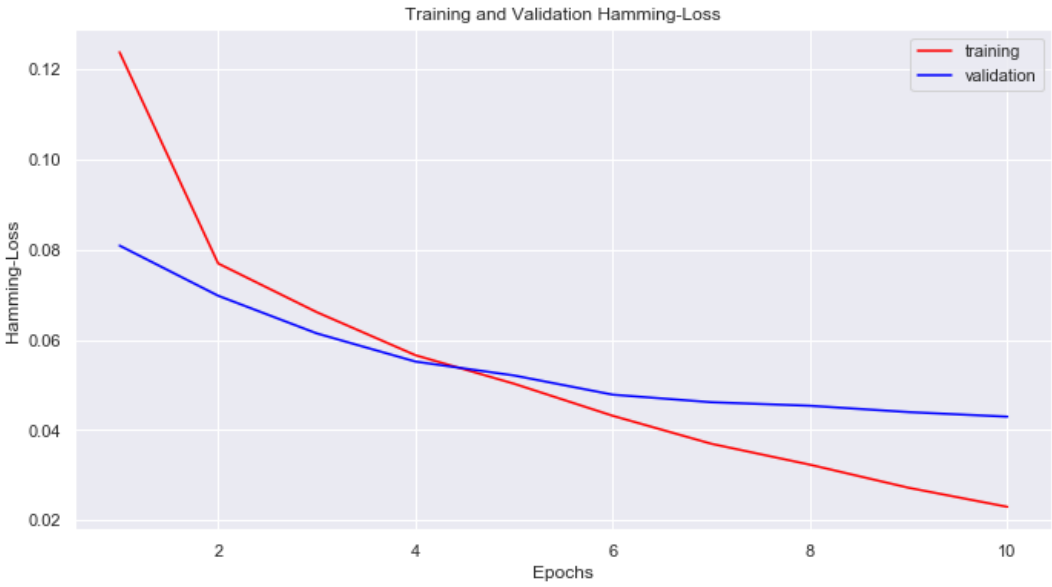
we can see that at 10th epoch there is a big difference in train and validation f1-micro-score and after 5th epoch training f1-micro is increasing faster as compared to validation f1-micro-score

Training and Validation F1-Macro-Score Plot



As we can see that after 6th epoch validation f1-macro-score is not increasing as much with respect to training f1-macro-score

Training and Validation Hamming Loss Plot



As we can se at epoch 5 train and validation are low and after that validaton loss is greater as compared to training

Conclusion

- Bert Large Uncased Model is performing slightly better as compared to Bert Base Uncased and also there is less overfitting till 10th epoch in Large model as compared to Base model