TensorBoard Observations

I preprocessed the `reviews` dataset as required, tokenized the text data, padded it to size 55, and generated the mask and segment vectors.

Used the pre-trained BERT model with 12 encoders and get the embedded vectors of our sentences by passing the padded sentence, mask arrays segment arrays into the BERT model.

Created a new NN with 4 Dense layers with 1024, 2048, 512, 1 neuron with the first 3 layers with `relu` activation function and last layer with 1 neuron is the output layer and `sigmoid` activation function. Used **BatchNormalization**, **Dropout** layers after each **Dense** layer to prevent the model from overfitting.

Compiled the model using the `adam` optimizer and `BinaryCrossentropy` loss function.

Used the `*Tensorboard*` callback to get logs and graphs of the model and training performance. Used the `*Checkpoint*` callback to save the best-performing weights. Used `*ReduceLROnPlateau*` to reduce the learning rate when a plateau condition is encountered.

Trained the model for 25 epochs and I got the best model at the 25th epoch and the `val_auc` would have improved if I would have trained for some more epochs.

So, I loaded the best-performed Model weights saved with Checkpoint callback and used them for predictions.

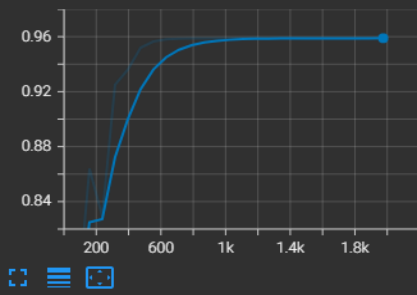I used this model to predict the class labels on the new data.

Note: Red is the train curve and blue is the validation curve.
1. The AUC is increasing over 25 epochs.
2. The loss is decreasing significantly after training for 25 epochs.
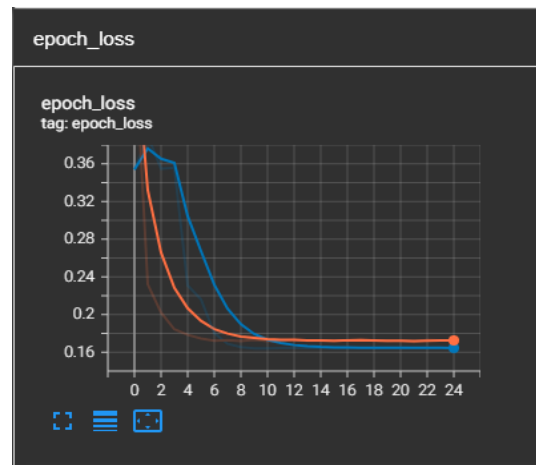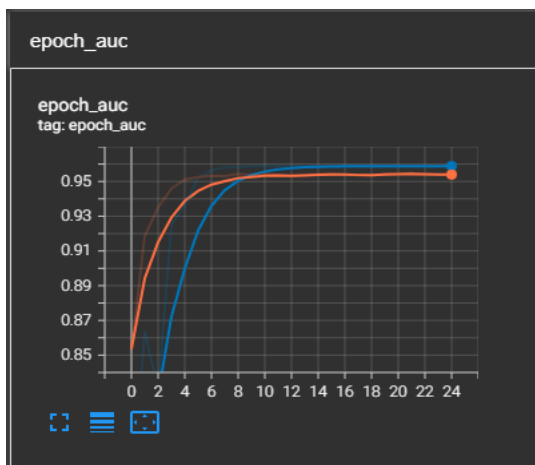
evaluation_auc_vs_iterations

evaluation_auc_vs_iterations
tag: evaluation_auc_vs_iterations

evaluation_loss_vs_iterations

evaluation_loss_vs_iterations
tag: evaluation_loss_vs_iterations