TensorBoard Observations

Reference - https://medium.com/neuronio/using-tensorboard-e3906a5798e6

Histogram observations:

On Tensorboard, they are used to visualize the weights over time. It is important because it could give a hint that the weights initialization or the learning rate are wrong.

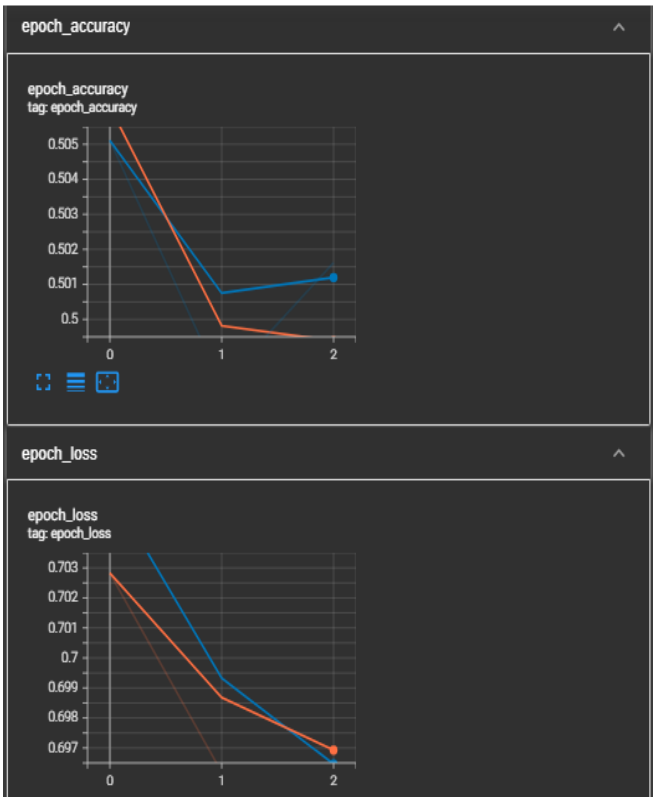
Due to random initializations, we get slightly different results every time. it is natural. For all the models, I used EarlyStopping Callback with min_delta = 0.35, the training stops at the 3rd epoch. The accuracy difference is less than 0.35.

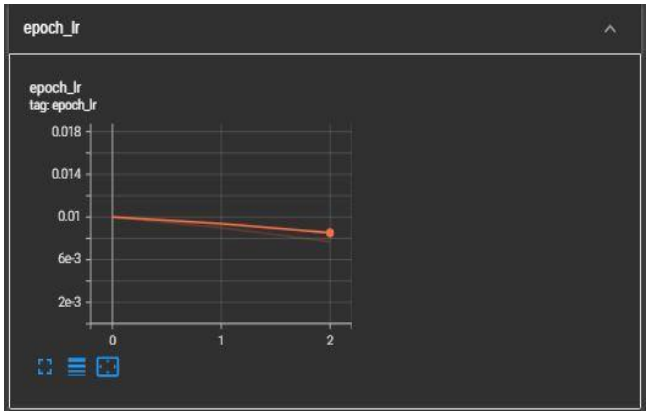
Model-1

In this model, the activation function is ***tanh*** for every layer except the output layer, ***SGD with momentum*** as an optimizer and the initializer is ***RandomUniform***.

Note: Red is the train curve and blue is the validation curve.

1. The epoch learning rate decreases from 0.01 to 0.007 from the first epoch to the third epoch.
2. The train epoch accuracy is 0.4966 and the validation epoch accuracy is 0.4986.
3. The train epoch loss is 0.6954 and the validation epoch loss is 0.695.





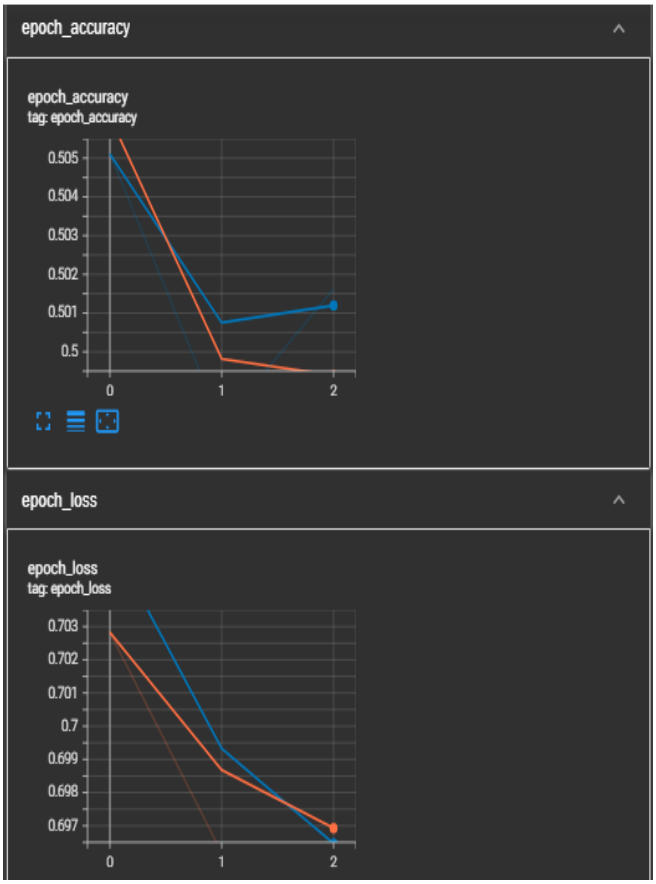
The micro F1 scores for each epoch are 0.5056060606060606, 0.49863636363636366, 0.5021212121212121

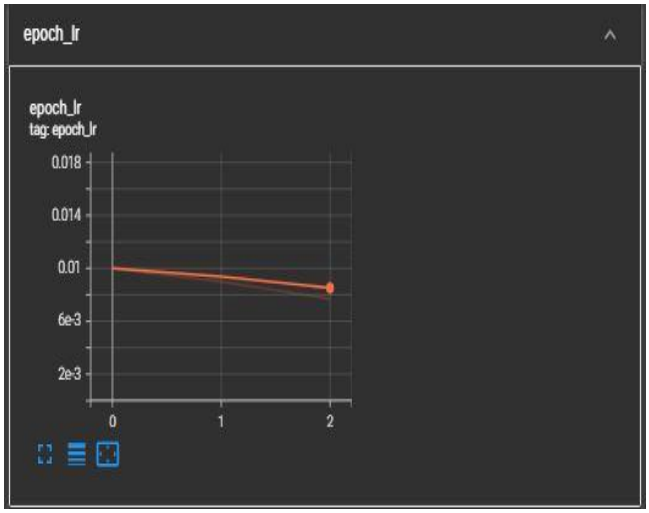
Model-2

In this model, the activation function is ***relu*** for every layer except the output layer ***SGD with momentum*** as an optimizer and the initializer is ***RandomUniform***.

Note: Red is the train curve and blue is the validation curve.

1. The epoch learning rate decreases from 0.01 to 0.007 from the first epoch to the third epoch.
2. The train epoch accuracy is 0.4966 and the validation epoch accuracy is 0.4986.
3. The train epoch loss is 0.6954 and the validation epoch loss is 0.695.





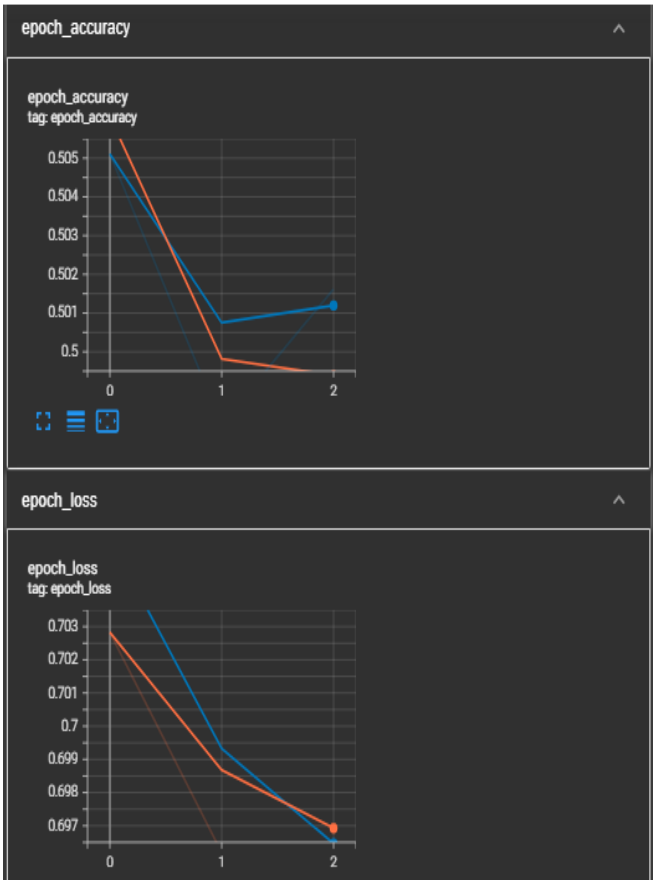
The micro F1 scores for each epoch are 0.5, 0.5, 0.5

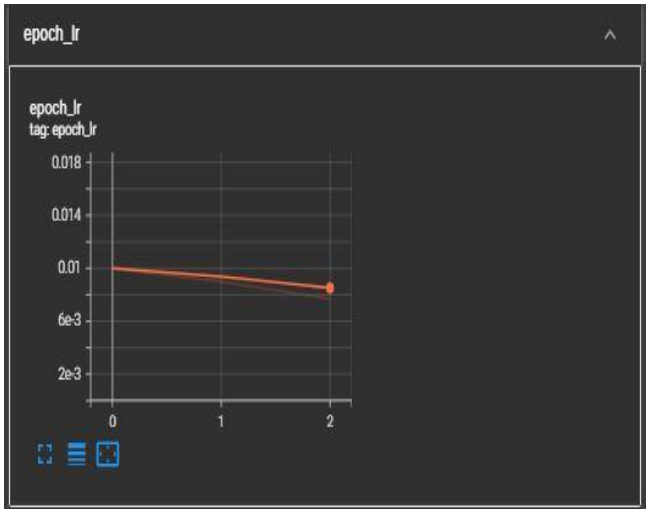
Model-3

In this model, the activation function is ***relu*** for every layer except the output layer, ***SGD with momentum*** as an optimizer and the initializer is ***HeUniform***.

Note: Red is the train curve and blue is the validation curve.

1. The epoch learning rate decreases from 0.01 to 0.007 from the first epoch to the third epoch.
2. The train epoch accuracy is 0.4966 and the validation epoch accuracy is 0.4986.
3. The train epoch loss is 0.6957 and the validation epoch loss is 0.695.





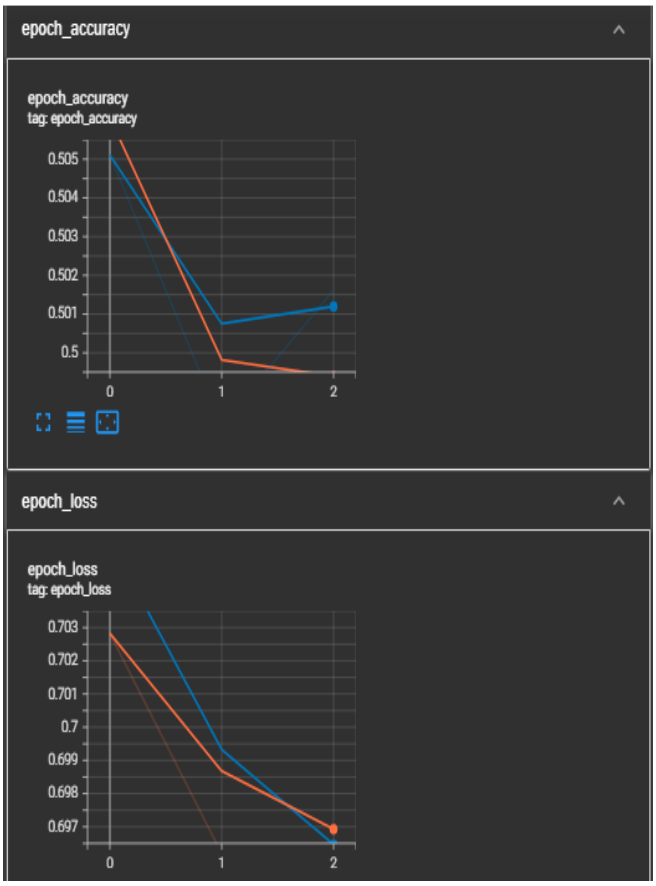
The micro F1 scores for each epoch are 0.6677272727272727, 0.6745454545454546, 0.6639393939393939

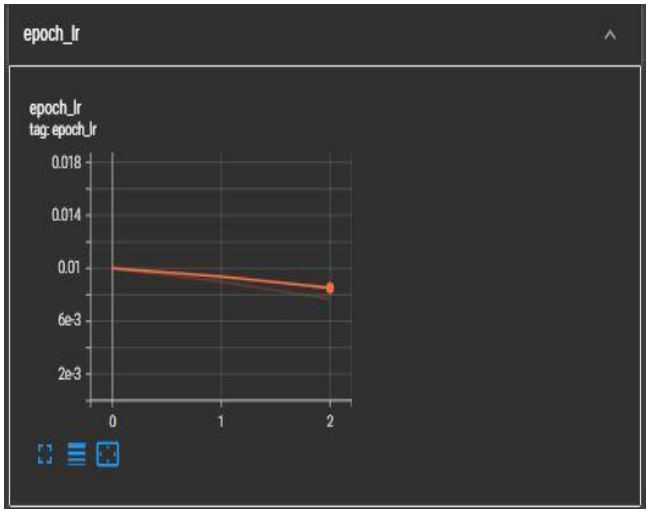
Model-4

Used ***relu*** as an activation for every layer except output layer, ***SGD with momentum*** as an optimizer, ***HeUniform()*** as initializer by changing the number of connections between the neurons in every layer.

Note: Red is the train curve and blue is the validation curve.

1. The epoch learning rate decreases from 0.01 to 0.007 from the first epoch to the third epoch.
2. The train epoch accuracy is 0.4966 and the validation epoch accuracy is 0.4986.
3. The train epoch loss is 0.6957 and the validation epoch loss is 0.695.





The micro F1 scores for each epoch are 0.5851515151515152, 0.6678787878787878, 0.6765151515151515