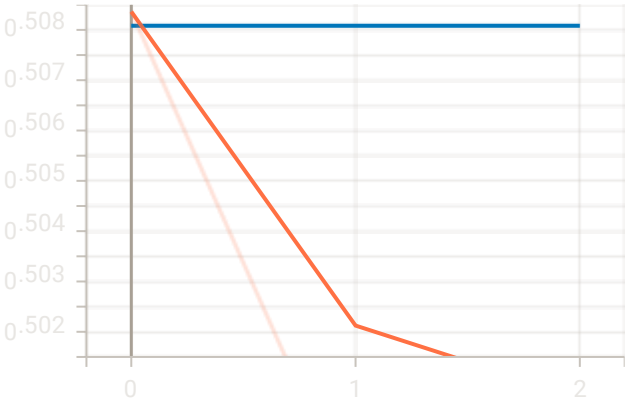
TensorBoard observations

1. Model 1:

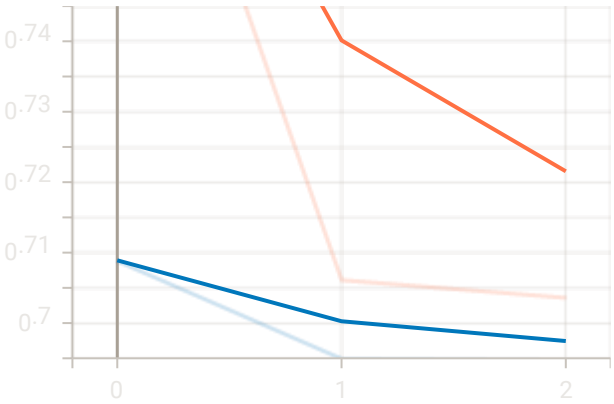
- 1. Use tanh as an activation for every layer except output layer.
- 2. use SGD with momentum as optimizer.
- 3. use RandomUniform(0,1) as initializer.
- 3. Analyze your output and training process.

- Epoch accuracy:



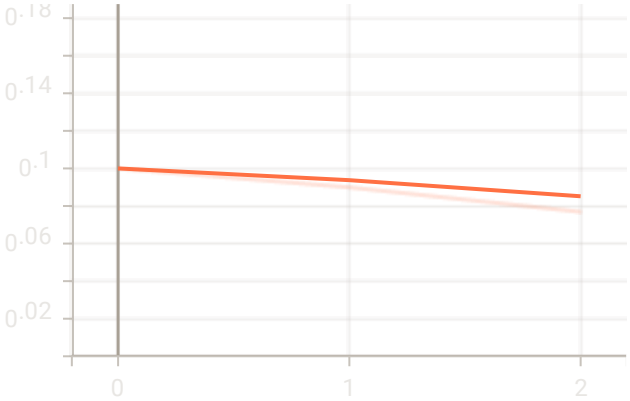
  - The validation accuracy remained the same throughout the epochs at about 0.50858.
  - However, the training accuracy decreased gradually down to a low of 0.4989.
  - The model ran until the third epoch where it stopped due to the EarlyStop callback

- Epoch loss:



  - The overall loss decreased to a total of 0.6948 on the validation data and 0.7215 on the training data.

- Epoch learning rate:



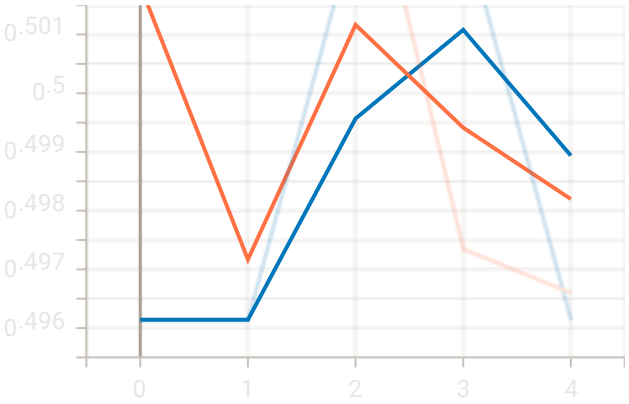
- The learning rate decreased from 0.1 to 0.08518 with an average decay of 10% on every epoch and an additional deduction of 5% on every third epoch.

Micro F1 – 0.491618, AUC – 0.494529

2. Model 2:

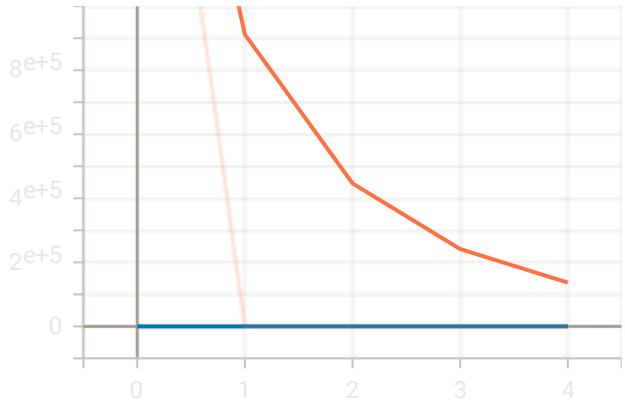
1. Use relu as an activation for every layer except output layer.
2. use SGD with momentum as optimizer.
3. use RandomUniform(0,1) as initializer.
4. Analyze your output and training process.

- Epoch accuracy:



- The validation accuracy reached a total of 0.50336 at the end of the fifth epoch.
- However, the training accuracy decreased gradually down to a low of 0.4987.
- The model ran until the fifth epoch where it stopped due to the EarlyStop callback

- Epoch loss:



- The overall loss decreased to a total of 0.6948 on the validation data and 0.1.3668e+5 on the training data.

- Epoch learning rate:

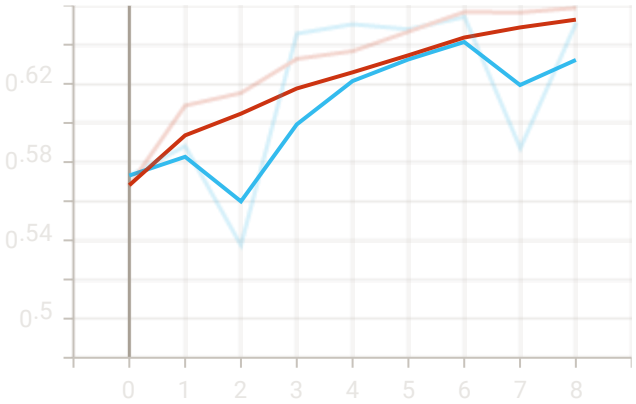
- The learning rate decreased from 0.1 to 0.0769499 with an average decay of 10% on every epoch and an additional deduction of 5% on every third epoch.

Micro F1 – 0.664238, AUC – 0.5

3. Model 3:

1. Use relu as an activation for every layer except output layer.
2. use SGD with momentum as optimizer.
3. use he_uniform() as initializer.
4. Analyze your output and training process

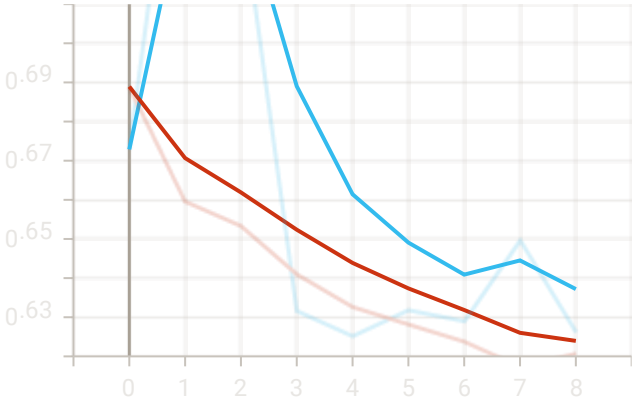
- Epoch accuracy:



- The validation accuracy reached a total of 0.6322 at the end of the ninth epoch.

- The training accuracy reached to 0.6528.
- The model ran until the ninth epoch where it stopped due to the EarlyStop callback, our longest run yet.

- Epoch loss:



- The overall loss decreased to a total of 0.6372 on the validation data and 0.624 on the training data.

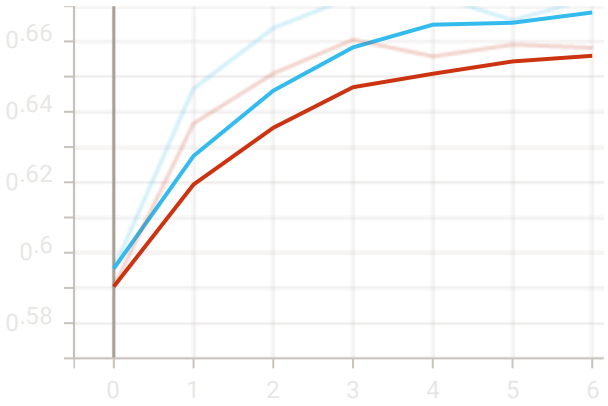
- Epoch learning rate:

- The learning rate decreased from 0.1 to 0.06250264 with an average decay of 10% on every epoch and an additional deduction of 5% on every third epoch. The learning rate hasn't decreased up to this low in our previous models.

Micro F1 – 0.430323, AUC – 0.611149

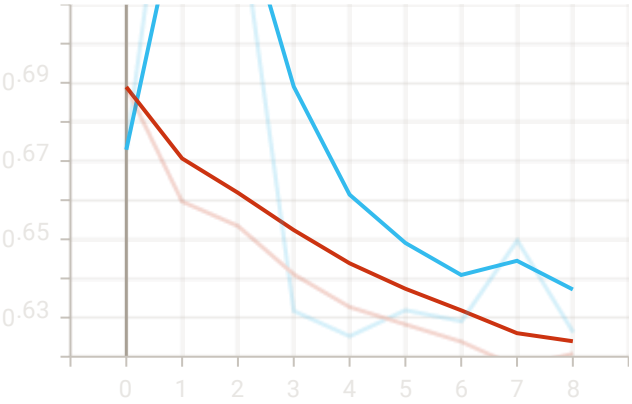
4. Model 4:

- Epoch accuracy:



- The validation accuracy reached a total of 0.6724 at the end of the seventh epoch.
- The training accuracy reached to 0.6582.
- The model ran until the seventh epoch where it stopped due to the EarlyStop callback.

- Epoch loss:



- The overall loss decreased to a total of 0.6056 on the validation data and 0.6084 on the training data.

- Epoch learning rate:
  - The learning rate decreased from 0.1 to 8.1225e-3 with an average decay of 10% on every epoch and an additional deduction of 5% on every third epoch. The learning rate hasn't decreased up to this low in our previous models. This low learning rate is the cause of using Adam optimizer.

Micro F1 – 0.675547, AUC – 0.6728912