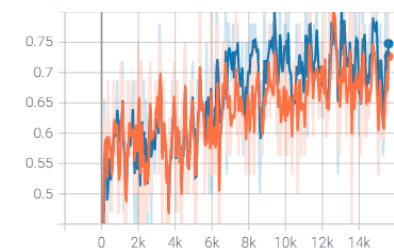


Tabnet 模型优化

epoch=10 的前提下，在 Attentive Transformer 分别使用 Sparsemax 和 Entmax-alpha（其中 $\alpha=1.5$ ）时：
训练集的表现

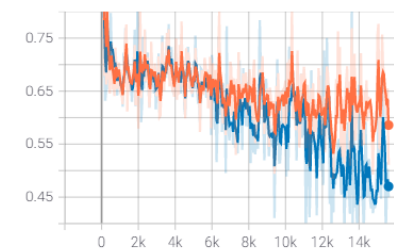
Training acc

Training acc



Training loss

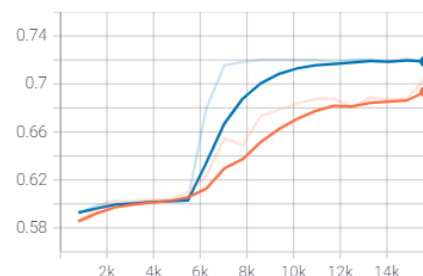
Training loss



验证集上的表现

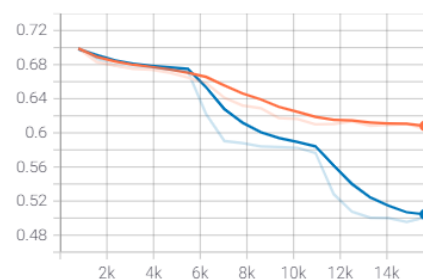
Validation acc

Validation acc



Validation loss

Validation loss



从上图可以看出，Entmax-1.5（蓝色曲线）在训练集和验证集的表现都明显优于 Sparsemax（橙色曲线）。