

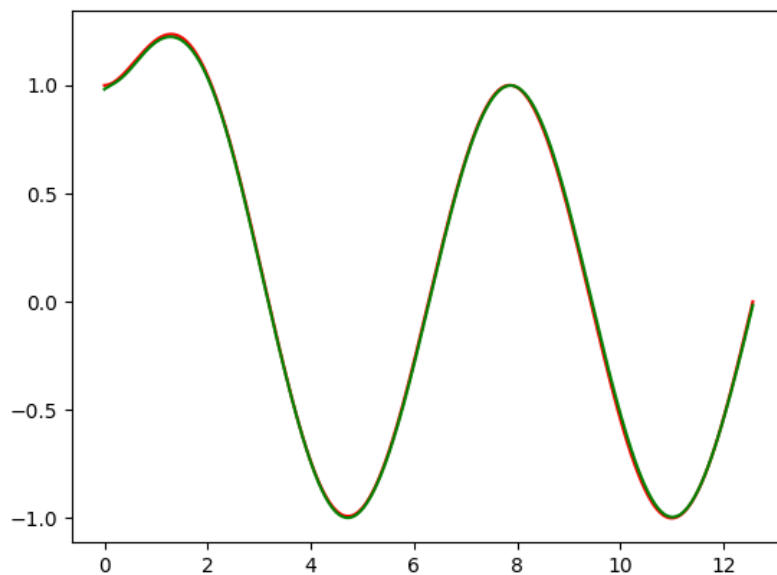
## 深度学习第一次实验报告

随机采样 10k 点作为训练集，8k 数据作为训练集，1k 数据作为验证集,1k 数据作为测试集。训练 100 个 epoch。

实验中最后选定的参数为：

adam-32-0.001-Tanh-3-32-0

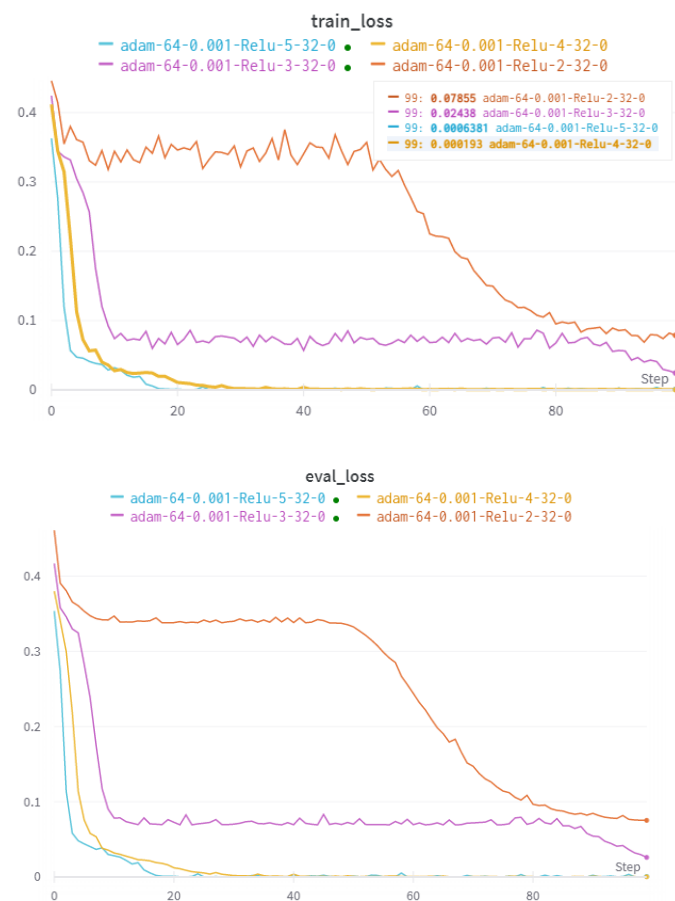
如下图为拟合结果，损失 = 0.000144863846716285



调参 1 改变网络深度 层数分别设置 2 3 4 5

(文件名 adam-64-0.001-Relu-3-32-0 中个参数为别对应：

优化器-bath size-学习率-激活函数-网络层数-隐层神经元个数-seed)

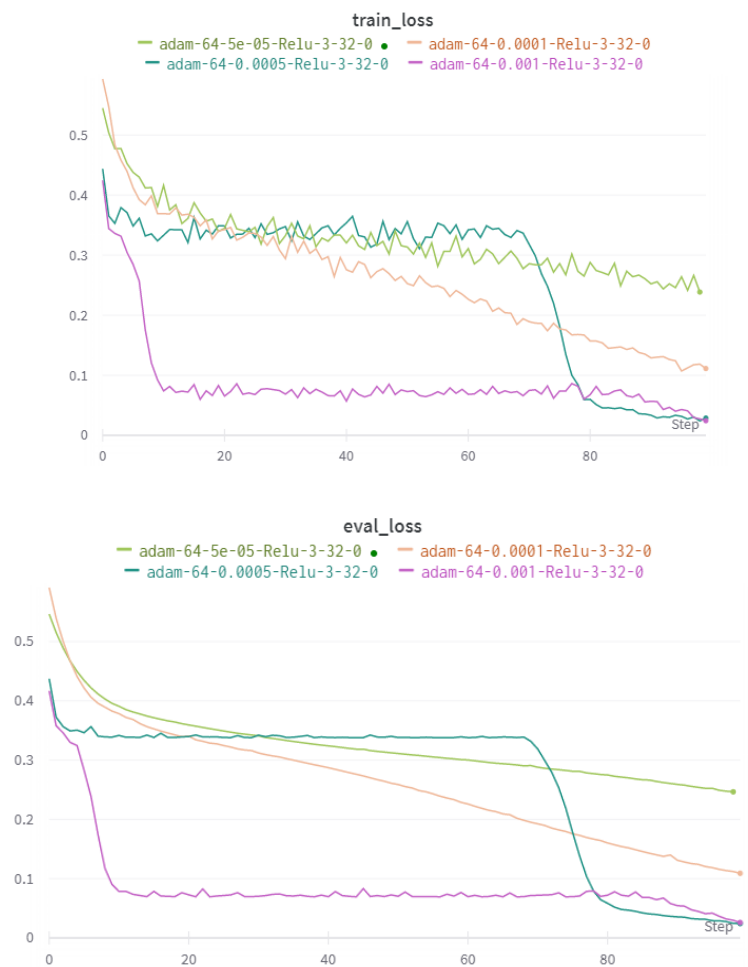


发现增大网络层数，网络的 loss 降低，网络拟合能力增强，收敛速度加快；当层数达到四层时，网络已经基本可以拟合函数；继续增大深度，性能几乎没有增加。

调参 2 改变学习率 学习率分别设为 1e-3 5e-4 1e-4 5e-5

(文件名 adam-64-0.001-Relu-3-32-0 中个参数为别对应：

优化器-bath size-学习率-激活函数-网络层数-隐层神经元个数-seed)

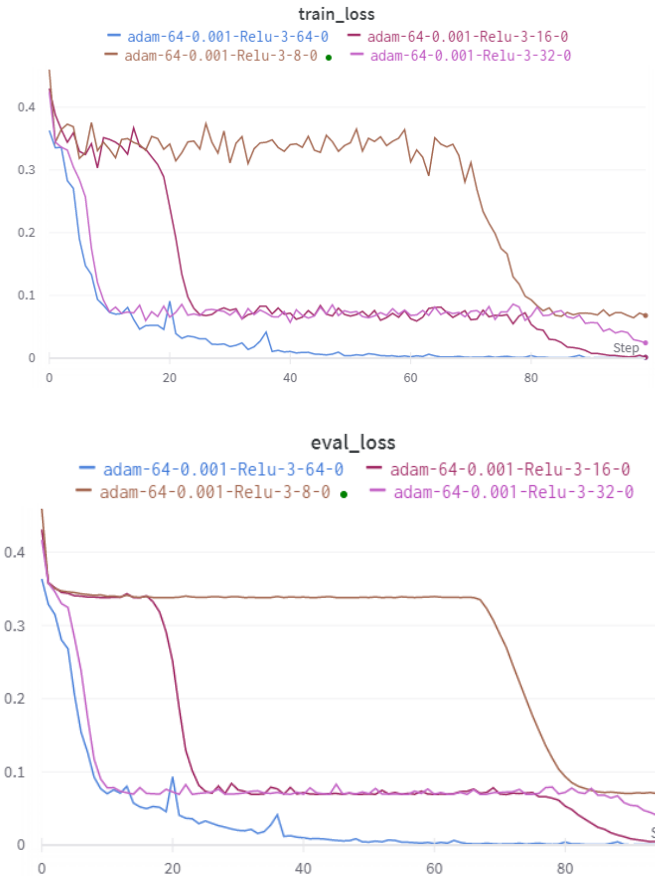


发现在学习率较大时，loss 会在几个 epoch 内迅速下降，且之后损失都比较小，而增大学习率，网络的损失缓慢下降，在 100 个 epoch 内没有达到收敛。

调参 3 改变网络宽度 隐层神经元个数分别设为 8 16 32 64

(文件名 adam-64-0.001-ReLu-3-32-0 中个参数为别对应：

优化器-bath size-学习率-激活函数-网络层数-隐层神经元个数-seed)



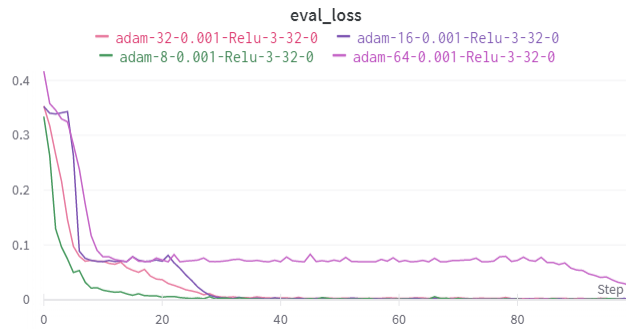
发现增大隐层神经元个数，网络收敛的 loss 降低。

调参 4 改变 batch size batch-size 分别设为 8 16 32 64

(文件名 adam-64-0.001-ReLu-3-32-0 中个参数为别对应：

优化器-bath size-学习率-激活函数-网络层数-隐层神经元个数-seed)



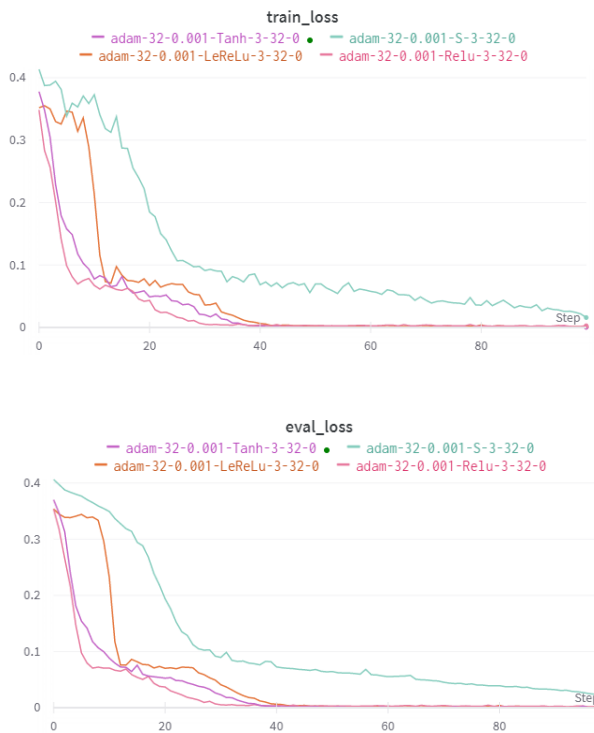


发现减小 batch size，使每个 epoch 内训练次数增多，网络收敛的更快而且收敛更低。

调参 5 改变激活函数 分别设为 ReLu LeakyReLu Sigmoid Tanh

(文件名 adam-64-0.001-ReLu-3-32-0 中个参数为别对应：

优化器-bath size-学习率-激活函数-网络层数-隐层神经元个数-seed)



从收敛结果和训练过程的稳定性考虑，最后选定的激活函数为 Tanh()