

- 1) 是否找到合适的损失函数? (不同问题适合不同的损失函数) (理解不同损失函数的适用场景)
- 2) batch size是否合适? batch size太大  $\rightarrow$  loss很快平稳, batch size太小  $\rightarrow$  loss会震荡 (理解mini-batch)
- 3) 是否选择了合适的激活函数? (各个激活函数的来源和差异)
- 4) 学习率, 学习率小收敛慢, 学习率大loss震荡 (怎么选取合适的学习率)
- 5) 是否选择了合适的优化算法? (比如adam) (理解不同优化算法的适用场景)
- 6) 是否过拟合? (深度学习拟合能力强, 容易过拟合) (理解过拟合的各个解决方案)
  - a. Early Stopping
  - b. Regularization (正则化)
  - c. Weight Decay (收缩权重)
  - d. Dropout (随机失活)
  - e. 调整网络结构