

陷入局部最优？

- sgd:调大学习率
- 换用adam等优化方法
- 递进学习->1000->10000->80000
- 调小batch: (A.全部数据作为一个batch和 B.每个采样作为一个batch) 一般用的时候, 在非监督训练阶段, 可以先用B方法创建初始模型, 然后在监督训练阶段, 选择小batch size进行初步训练, 让模型跳出局部极值, 之后用大的batch size让模型收敛, 这样一般能达到比较好的效果。