1. 验证集和测试集应该\_\_\_分布？ （同一/不同）
2. 神经网络模型高方差或高偏差时，分别有哪些可尝试的策略？
3. 梯度消失或爆炸产生的原因及解决方法？