1. 验证集和测试集应该\_\_同一\_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. 正则化-加入正则项，droup out，数据增强，early stoping --过拟合（高方差）
4. 调整损失函数和网络结构---欠拟合（高偏差）
5. 梯度消失或爆炸产生的原因及解决方法？

原因：

梯度消失产生的原因：

（1）隐藏层的层数过多;【网络层之间的梯度（值大于 1 or小于1）重复相乘导致的指数级增长/下降会产生梯度爆炸/梯度消失。】

（2）采用了不合适的激活函数。

梯度爆炸产生的原因：

（1）隐藏层的层数太多;

（2）权重初始化值过大。

解决方法：

1. 合理的初始化权重
2. 选择合理的激活函数（在神经网络中relu比sigmoid[f’(x) = f(x)(1-f(x)),两个小于1的数相乘更小]更合适）
3. 残差结构
4. Batch Normalization-快且调节参数w,b
5. 梯度截断（设置一个上限）