1. 验证集和测试集应该同一分布？ （同一/不同）

为了定位问题，验证集要用来评估模型的效果

1. 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？

* 高方差：数据过拟合

1. 新网络，比如：更多的隐藏层或隐藏单元
2. 用更多时间训练算法
3. 更先进的优化算法
4. 尝试新网络框架（不一定有用，一般更大的网络拟合性能更好）
5. early stopping：提前结束优化损失函数
6. 添加多项式项（高次）

* 高偏差：数据欠拟合

1. 增加数据量
2. 正则化
3. 尝试新网络框架（不一定有用，有可能一箭双雕）
4. 梯度消失或爆炸产生的原因及解决方法？

* 梯度消失

深层神经网络中权重小于1

* 梯度爆炸

深层神经网络中权重大于1

**解决方法：**

* 随机初始化

对于ReLu: *Herd et al.*

parameters['W' + str(l)] =

np.random

.randn(layer\_dims[l], layer\_dims[l-1]) \*

(2/np.sqrt(layer\_dims[l-1]))

对于tanh：常数1

* 使用ReLu激活函数

ReLu：非饱和的激活函数

* normalization规范化（batch normalization）
* 梯度截断：缓解梯度爆炸（if d>a: d=a）