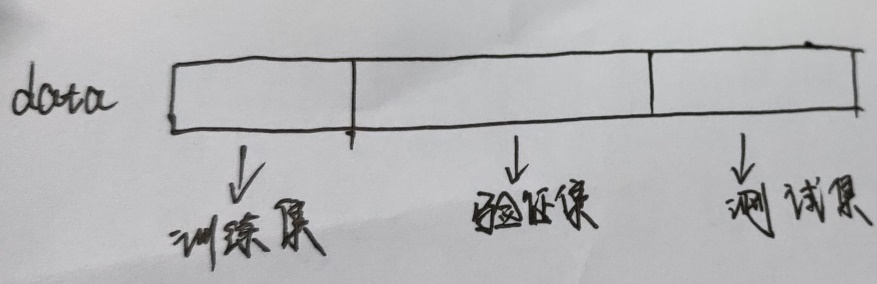
**改善深层神经网络：超参数调试、正则化以及优化**

第一周：深度学习的实用层面

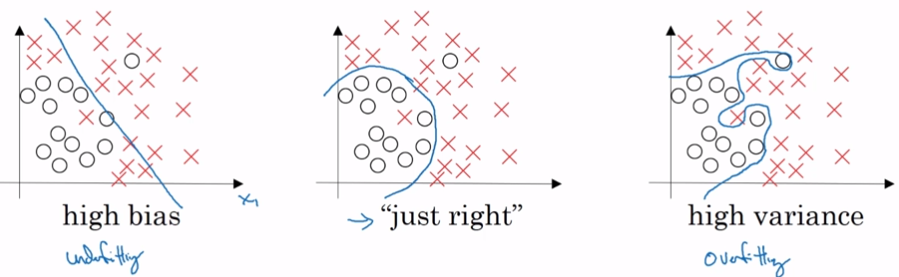
1.1 训练开发测试

假设这是一个训练数据，用长方形表示，如图：通常将这些数据划分为几部分，分别为训练集、验证集、测试集，接下来开始地训练集执行训练算法经过验证集选出最好的模型，经过充分验证，选择出最终的模型然后就在测试集上进行评估。要确保验证集和数据集来自同一分布。测试集的目的是对：最终所选定的神经网络系统做出无偏评估。

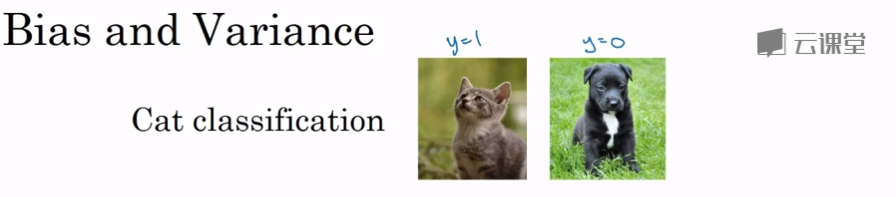


1.2偏差和方差

如果给这个数据集拟合一条直线，可能得到一个逻辑回归拟合，这是偏差高的情况，称为“欠拟合”，相反，如果拟合一个非常复杂的分类器，比如深度神经网络但也不是一个好的分类器，分类器的方差较高，数据过度拟合。还有像中间图那样的复杂程度适中，数据拟合适度的分类器，称为适度拟合。



以猫咪图片分类例子：



训练集误差：1% 15% 15% 0.5%

验证集误差：11% 16% 30% 1%

假定训练集错误率为1%，假定验证集错误率为11%，可以看出可能过渡拟合了训练集，某种程度上，验证集没有充分利用交叉验证集的作用，这种称为“高方差”。衡量训练集和验证集的误差，得出不同的结论；假设训练集误差为15%，验证集误差为16%，假设该案例中人的错误率为0%，算法没有在训练集中得到很好的训练，如果训练的拟合度不高，就是“欠拟合”。可以说这种算法偏差较高。相反，验证集产生的结果确是合理的。

最优误差也被称为“贝叶斯误差”。

通过看训练集的误差，可以判断数据的拟合情况，可以判断是否又有偏差的问题，然后查看错误率有多高；当完成训练，开始验证时，可以判断方差是否过高。

1.3机器学习基础

初始化模型训练完成之后，我们要先知道算法的偏差高不高，如果偏差较高，试着评估训练集或者训练数据的性能， 如果偏差的确很高，甚至无法拟合数据，那么应该选这一个新的网络，花费更长时间来训练算法，尝试更加先进的算法。训练学习算法时，我们会不断的尝试这些方法，直到解决到偏差问题反复尝试，直到拟合数据为止。一旦方差降低到一个可接受的值，检查下方差有没有问题，为了评估方差，我们要查看验证集的性能，如果方差过高，最好的解决方法就是采用更多的数据，也可以 通过利用正则化来减少过拟合。总之就是不断的重复尝试，直到找到一个底偏差、底方差的框架。

1.4正则化

解决高方差的问题：1、正则化 2、用更多的数据

正则化的原理：

求成本函数的最小值，在后面加入一个λ，（正则化参数）

L2正则化：直接在原来的损失函数基础上加上权重参数的绝对值。

