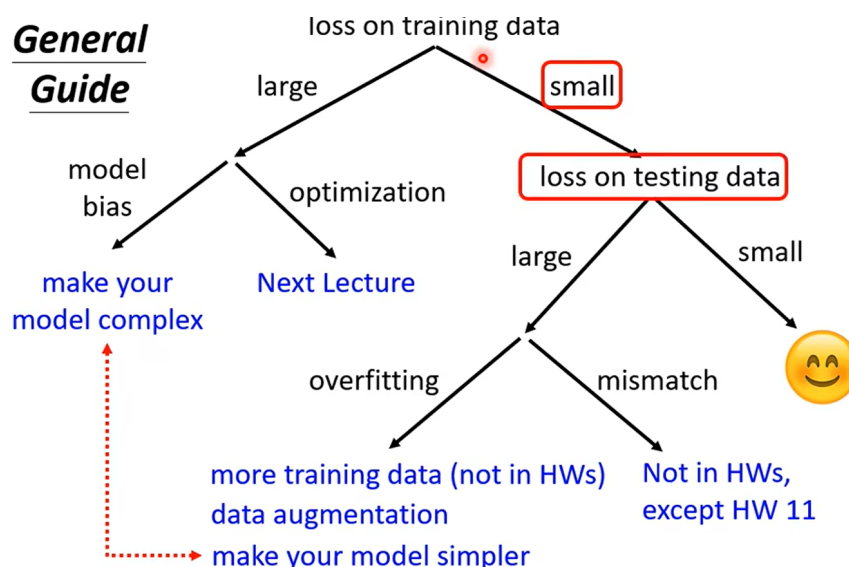


Week2

当模型效果较差时，调节模型的思路。



▼ loss on training data较大时：

1.模型不够复杂

增加模型的深度或者feature的数量去尝试。

2.模型优化不到位

更简单的模型更加容易优化到最优解，因此可以先尝试采用简单的模型去优化，再采用比较复杂的模型。

▼ 当loss on training data 较小时，在训练集上得到了比较好的结果：

1.若loss on testing data比较大，那么此时就发生了过拟合。

相对于样本来讲，模型的复杂度过高。

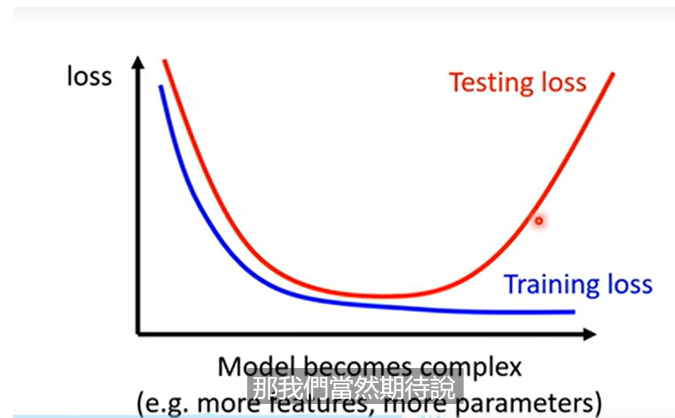
▼ 解决方案有如下几种：

1.可以增加样本的数量，一般用不到。

2.可以进行data augmentation，即数据增强技术，

3.可以尝试对模型进行限制。限制策略包括减少参数的数量（或者让参数间有关），或者减少输入的feature数目。也可以early stopping, regularization和

dropout。



模型复杂度和loss之间的关系，复杂度过高的模型容易出现过拟合现象。

▼ 当loss on training data和loss on testing data都比较小时，有可能会出现问题：

1.测试结果好的模型不一定有好的效果，有可能只是输出比较适合测试而已。

cross validation可以将训练数据的一部分不训练，作为validation，从而挑选模型，然后再在testing data上进行测试，一定程度上可以避免这种情况。

N-fold Cross Validation方法：

Training Set			Model 1	Model 2	Model 3
Train	Train	Val	mse = 0.2	mse = 0.4	mse = 0.4
Train	Val	Train	mse = 0.4	mse = 0.5	mse = 0.5
Val	Train	Train	mse = 0.3	mse = 0.6	mse = 0.3
Avg mse = 0.3			Avg mse = 0.3	Avg mse = 0.5	Avg mse = 0.4

N-fold cross validation方法验证模型

2.mismatch，指的是training data和testing data的分布不一样而产生的问题，解法待定。