



1. High bias解决方案: Boosting、复杂模型（非线性模型、增加神经网络中的层）、更多特征

2. High Variance解决方案: bagging、简化模型、降维

具体而言

高偏差，可以用boosting模型，对预测残差进行优化，直接降低了偏差。也可以用高模型容量的复杂模型（比如非线性模型，深度神经网络），更多的特征，来增加对样本的拟合度。

高方差，一般使用平均值法，比如bagging，或者模型简化/降维方法，来降低方差。

高偏差和高方差都是不好的，我们应该加以避免。但是它们又是此消彼长的关系，所以必须权衡考虑。一般情况下，交叉验证训练可

以取得比较好的平衡：

将原始样本均分成K组，将每组样本分别做一次验证集，其余的K-1组子集数据作为训练集，这样会得到K个模型，这K个模型可以并发训练以加速。用这K个模型最终的验证集的分类准确率的平均数作为此K-CV下分类器的性能指标。K一般大于等于3，而K-CV 的实验共需要建立 k 个models，并计算 k 次 test sets 的平均预测正确率。

在实作上，k 要够大才能使各回合中的 训练样本数够多，一般而言 k=10（作为一个经验参数）算是相当足够了。