

## 第十次作业.

黄磊 2022E2013282156.

1. 模型复杂度过低, 易欠拟合, 引起偏差大;  
模型复杂度过高, 易过拟合, 引起大方差;

2. 判断: 通过训练集和验证集准确率判断

缓解: 过拟合: 提前结束训练  
减少特征数量  
提高正则化水平  
扩大训练集

欠拟合: 增加迭代次数  
更多特征  
降低正则化水平.

3. Bagging: 在数据集上  $D$  进行 Bootstrap 采样, 后训练模型.  
重复  $M$  次, 得到  $M$  个模型, 采用其平均投票  
$$f_{\text{avg}}(x) = \frac{1}{M} \sum_{i=1}^M f_i(x)$$

Boosting: 按序列学习, 弱学习器组合成强学习器.

4. AdaBoosting:

对于一系列弱学习器训练, 对于  $\phi_k$  中错误的样本,  
在  $\phi_{k+1}$  中权重更大, 正确分类样本权重减小.

5. 17: A B. B

