

总结过拟合与欠拟合的原因及解决办法

笔记本： 我的第一个笔记本

创建时间： 19/10/27 12:46

更新时间： 19/10/27 13:03

作者： 2113751293@qq.com

总结过拟合与欠拟合的原因及解决办法

过拟合

- 原因
 - 数据过于复杂
 - 数据太少
 - 数据不均衡
- 解决办法
 - 降低模型复杂度
 - 使用正则化方法(L1, L2, Dropout)
 - 对数据进行上采样或下采样得到均衡数据
 - 增大数据量
 - 重新清洗数据, 提出异常值
 - 早停
 - 交叉验证
 - 剪枝
 - 集成学习(boosting, bagging)

欠拟合

- 原因
 - 模型过于简单
 - 训练时间不够
 - 使用的特征与目标任务关系不大
- 解决办法
 - 使用更复杂的模型
 - 训练足够的时间
 - 寻找更为关键的特征
 - 减弱正则化效果

Writing down three sceneries that machine learning has been used now

- 银行反欺诈
- 电子客服

Come out with three new sceneries with which machine learning may be applied.

- 同声传译
- 人工智能芯片