

ML概述_AI07_20190424

一、多选题

1、机器学习是构建一个映射函数，应用场景有哪些（ ）

- A、语音识别
 - B、图像识别
 - C、围棋
 - D、对话系统
-

2、ROC曲线的横、纵坐标分别表示？（ ）

- A、TNR, TPR
 - B、FNR, FPR
 - C、FPR, TPR
 - D、FPR, FNR
-

3、属于机器学习问题的有哪些？（ ）

- A、分类
 - B、回归
 - C、聚类
 - D、降维
-

4、ML的三要素有哪些（ ）

- A、模型
 - B、学习准则
 - C、优化
 - D、迭代
-

5、模型选择方法有哪些？（ ）

- A、留出法

B、交叉验证法

C、自助法

D、过滤法

6、模型评估的方法包括哪些 ()

A、Confusion Matrix

B、ROC

C、AUC

D、Lift

E、Gain

7、以下说法正确的是 ()

A、过拟合：训练误差和泛化误差的差距太大。

B、欠拟合：是指模型不能再训练集上获得足够小的误差 (训练误差太大)

C、泛化错误：一般表现为一个模型在训练集和测试集上错误率的。可以衡量一个ML模型是否可以很好地泛化到未知数据。

D、ML目的：减少泛化误差。

8、当模型欠拟合或者过拟合了，以下哪种说法是对的？ ()

A、欠拟合处理：增加模型复杂度

B、过拟合处理：降低模型复杂度

C、欠拟合处理：正则化处理

D、过拟合处理：增加数据集

(圈里说法；好的数据集比好的模型更重要。)

二、简答题

1、判断一个模型是否过拟合的依据是什么？

2、什么是ML？

3、留出法交叉验证 (holdout cross-validation) 整个过程描述。

4、期望误差是越小越好？

5、超参数和学习参数的区别

6、留出法交叉验证 (holdout cross-validation) 和k-fold交叉验证的优缺点。

7、Bootstrap在什么场合上用到，采样过程是什么样子？

8、写出Confusion Matrix模型评估中准确率 (Accuracy)、精准率 (Precision)、召回率 (Recall)、F值的表达式。

真实情况

预测结果

预测结果

正例

正例

反例

TP

FN

反例

FP

TN

9、ROC的全名是什么？ROC和AUC曲线横纵坐标分别表示什么？
