

2021年机器学习_期末真题_回忆版

【author】Maxpicca-Li

【写在前面】

1. 考试时间：2021年12月28日16:30~18:30。
2. 试卷整体简单，题量也非常少，但是我的命中率不太行，呜呜呜。
3. 本回忆版真题于2021年12月28日18点01分写成，因为朴素贝叶斯emo了（咳咳，郑重申明，没有泄题，我已经出来了才写的）。
4. 计院专业课的试卷似乎都不准老师发出来，希望有学弟学妹们能将我“回忆试卷”的习惯传承下去！
——from VayneDuan
5. 其余专业课的回忆版试卷如下：记得 star & follow, 会持续更新的！

Source: [VayneDuan/cqu-cs-learning-materials](https://github.com/VayneDuan/cqu-cs-learning-materials)

Fork: [Maxpicca-Li/cqu-cs-learning-materials](https://github.com/Maxpicca-Li/cqu-cs-learning-materials)

——【写在前面】基于原作者VayneDuan修改

简答题（5道*10分=50分）

1. 简述假设空间和版本空间。
2. 简述训练误差、测试误差、泛化误差。训练误差很小，测试误差一定很小吗？
3. 对数几率回归解决的是回归问题还是分类问题？主要通过什么方法训练样本？
4. 朴素贝叶斯主要解决了什么障碍？其关键假设是什么？
5. 简述局部极小和全局极小。

算法题（2道*10分=20分）

1. 为k-means算法的伪代码添加注释。
2. 补全决策树的伪代码。

综合题（2道*15分=30分）

1. 支持向量机基本型解决的是回归问题还是二分类问题？如果超平面无法在训练集样本中进行划分，请问还可以用支持向量机基本型吗？如果不可以用，请问可以使用什么改进方法？改进的基本原理是什么？
2. 集成学习的性能一定会提升吗？集成学习的关键影响因素是什么？是如何影响的呢？请枚举一种可以提高集成学习的性能的方法。