

2024 统计机器学习试卷（回忆版）

简答题（题目顺序可能有些差错）

- 第一题：什么是机器学习；监督学习和无监督学习是什么，分别举一个具体例子
- 第二题：什么是过拟合现象，列出防止过拟合的方法（至少三种）
- 第三题：knn和kmeans的区别
- 第四题：什么是核函数，常见的核函数有哪些（至少三种）
- 第五题：交叉验证的作用是什么，描述至少两种交叉验证方法
- 第六题：机器学习涉及的隐私和安全问题有哪些？可以采取哪些措施？

计算题（题目给出了模型描述）

- 第一题：kmeans聚类
- 第二题：朴素贝叶斯（不带拉普拉斯平滑）
- 第三题：决策树（难度常规，但计算量较大，建议熟练掌握）