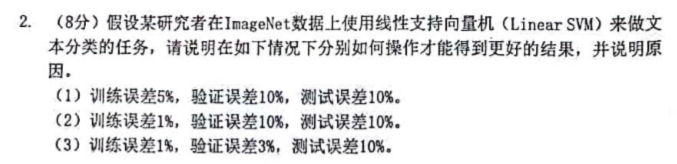
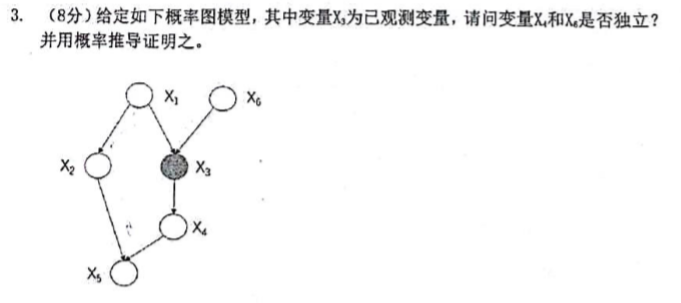


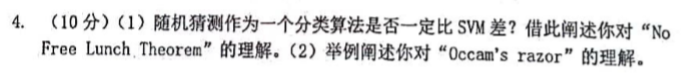
1. 广义的说，存在于时间和空间中可观测的物体，如果我们可以区别它们是否相同或者相似，都可以称之为模式。模式所指的不是事物本身，而是从事物获得的信息。因此模式往往指的是具有时间或空间分布的信息。
2. 模式的直观特征：可观察性，可区分性，相似性
3. 主要方法：
   1. 监督学习：概念驱动，归纳假说。
   2. 非监督学习： 数据驱动，演绎假说。



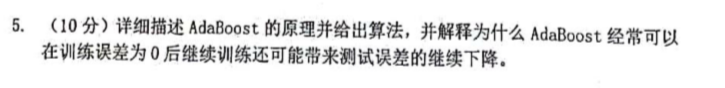
1. 欠拟合，适当的增大C值，减少错分样本。
2. 过拟合，适当的降低C值，增加模型的泛化能力。
3. 训练数据和测试数据不是独立同分布，建议重新采样或者shuffle数据。



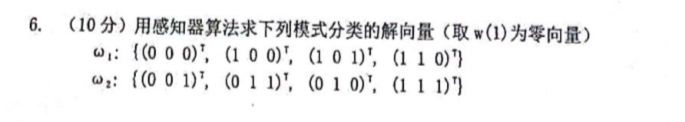
得证已知的情况下，和独立。



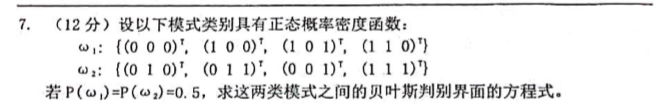
1. 脱离具体问题谈论算法优劣是没有意义的，在特定的问题上随机猜想是可以比SVM好的。
2. No Free Lunch Theorem : 在问题等概率出现且等权重的情况下，任何算法的期望都是一样的。也就是说，没有一个算法可以在任何问题上总是产生最好的分类器。脱离具体问题讨论算法的优劣是无意义的。只有针对具体问题的具体模型，才能对比优劣。
3. Occam’s razor ：这是一种归纳偏好： 如无必要，勿增实体。达到相近性能的模型中，最简单的往往更加接近真相。过度复杂只会造成过拟合而失去泛化能力。



1. AdaBoost 原理： 基于强分类器比较难以获取，期望训练多个弱分类器配合构成一个强分类器的思想，AdaBoost使用在弱分类器1上训练失败的样本去训练弱分类器2的思路，通过调整样本权重，使得弱分类器1在样本上等价于随即猜想，然后用调整后的权重样本去训练分类器2.
2. AdaBoost 算法：
   1. 初始化样本权重
   2. 迭代
      1. 在权重下训练弱分类器 ，要求分类器性能优于随即猜想。
      2. 计算
      3. 更新权重因子
         1. 其中 ; 是归一化因子。
   3. 最终的训练器是
3. 训练误差为0后AdaBoost继续训练类似于继续寻找更大的分类margin



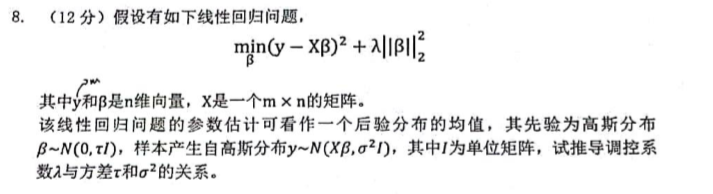
1. 获得规范增广矩阵
2. 初始化
3. 迭代
   1. 第一轮全军覆没
   2. 第二轮 错误
   3. 第三轮第二列错误
   4. 第四轮第一列错误
   5. 第五轮 错误，
   6. 第六轮错误
   7. 第七轮全部完成，解向量



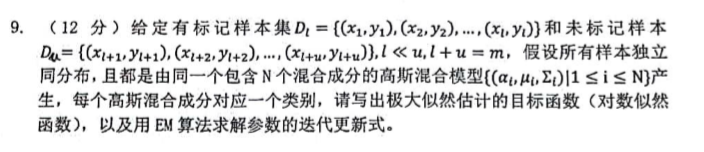
可见，又由于 , 因此这是一个最小马氏距离分类器。

分类界面方程

就不继续计算了！give up



因此，最大似然等价于 ,因此

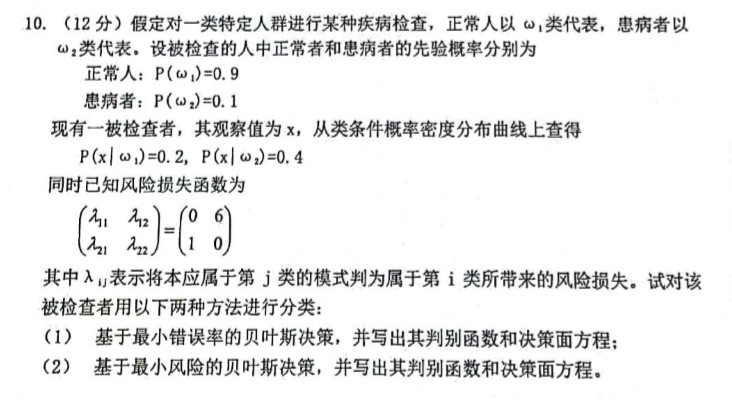


1. 极大似然估计的目标函数

令

1. EM 算法求参数的迭代方式
   1. 初始化一个
   2. 迭代
      1. 根据当前求的分布（无标签部分）
      2. 有标签部分
      3. 利用的最大似然估计，用估计值更新

参数迭代公式（背过吧~）



1. 最小错误率贝叶斯：
   1. 判别函数 ：
   2. 决策面方程
   3. 决策

判决属于。

1. 最小风险贝叶斯：
   1. 判别函数

* 1. 决策面方程
  2. 决策：

故判决属于