

第六讲 课后作业

两类分类问题：

- 1) 文件 Ionosphere Dataset.csv 中包含 351 个样本，前 200 个样本组成训练样本集，后 151 个样本组成测试样本集。每个样本由 34 个特征描述，两个类别分别使用 g 和 b 来标记。
- 2) (30)编程实现线性判别式最小二乘法算法。使用训练样本得到线性判别式模型，分别给出训练样本错误率和测试样本错误率。
- 3) (30)编程实现 Fisher 线性判别式算法。使用训练样本得到线性判别式模型，分别给出训练样本错误率和测试样本错误率。
- 4) (30)编程实现概率生成式模型（假设类条件概率密度为正态分布）。使用训练样本得到概率生成式模型，分别给出训练样本错误率和测试样本错误率。
- 5) (10)比较 2)和 3)的结果，试解释原因。