

多元正态分布下的贝叶斯决策

于虎

邮箱: huyu@mail.bnu.edu.cn

课程指导教师: 郭小娟 教授
北京师范大学人工智能学院

提纲

- 实验背景与目的
- 实验要求
- 实验步骤
- 实验内容
- 实验结果分析
- 实验提交

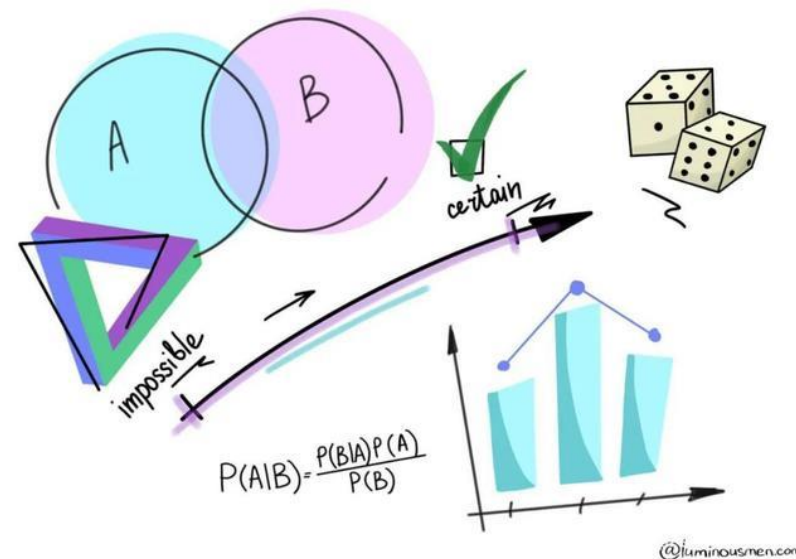
实验背景与目的

• 实验背景：

在模式识别领域，贝叶斯决策是一种基于概率统计的决策方法，广泛应用于图像识别、语音识别、机器学习等领域。贝叶斯决策的核心思想是利用先验概率和条件概率，在给定观测数据的情况下，通过计算后验概率来进行决策。多元正态分布是贝叶斯决策中常用的概率模型，它假设数据服从多元正态分布，通过建立模型的参数，可以进行分类和识别。

• 实验目标：

- ① 了解贝叶斯决策的基本原理和应用场景；
- ② 理解多元正态分布的概念和特性；
- ③ 掌握使用贝叶斯决策进行模式识别和分类的方法；
- ④ 学会使用实验工具和技巧，对模型进行训练和测试；
- ⑤ 分析和评估实验结果，提出改进和扩展的可能性。



实验要求

- ① 理解两类贝叶斯决策的基本概念；
- ② 熟练掌握正态分布概率模型下的最小错误率贝叶斯决策的过程；
- ③ 要求写出每个实验对应的算法原理并写出对应代码的测试步骤；
- ④ 本实验除画图可以调用对应包，其它部分禁止调包。

实验步骤

- ① 理解贝叶斯决策算法的原理和基本步骤。
- ② 了解多元正态分布的概念和性质，以及如何生成符合多元正态分布的随机样本。
- ③ 设计实验数据集，包括特征向量和类别标签。
- ④ 划分数据集为训练集和测试集，通常采用交叉验证方法。
- ⑤ 根据训练集数据，计算各类别的先验概率和条件概率密度函数。
- ⑥ 利用贝叶斯决策算法进行分类，将测试集的样本分到概率最大的类别中。
- ⑦ 计算分类结果的准确率、精确率、召回率和F1值等评价指标。
- ⑧ 重复实验步骤5至7，对不同的训练集和测试集进行分类。
- ⑨ 比较不同数据集上贝叶斯决策算法的分类性能，并与其他分类算法进行比较。
- ⑩ 分析实验结果，探讨贝叶斯决策算法的优缺点和适用范围。
- ⑪ 撰写实验报告，详细描述实验步骤、实验结果和分析、实验结论等内容。

实验内容

1、问题描述

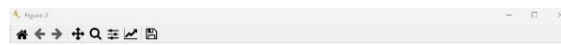
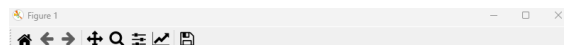
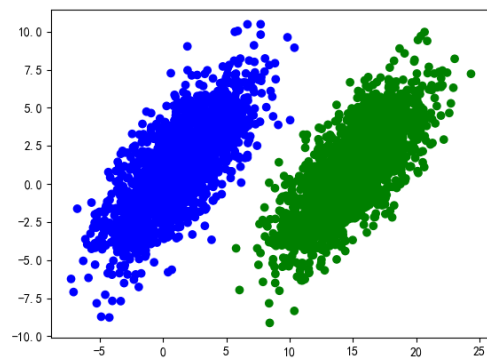
分别生成两类数据集，每类数据集的样本是3000个，有2个特征，服从不同的二维高斯分布。对其随机划分成训练集（4000个样本）和测试集（2000个样本），在不同的前提下分别进行贝叶斯决策。

- ① 每类的协方差矩阵相等，且2个特征相互独立（对角线矩阵），类间均值不同。
- ② 每类的协方差矩阵相等，且2个特征不相互独立（对角线矩阵），类间均值不同。
- ③ 每类的协方差矩阵不相等，类间均值不同。

实验结果分析

- ① 画出两个不同高斯分布的等概率密度线以及概率密度图；
- ② 计算测试集的准确率，并画出决策面；

```
Run main x
D:\Anaconda\envs\SHYBOYXH\python.exe
分类准确率: 0.9995
Process finished with exit code 0
```



实验提交

- 模式识别-实验一 统计决策方法
- 链接地址：<https://pan.bnu.edu.cn/l/h1QPVB>
- 命名方式：
 - ① “学号_姓名”
 - ② 需要有“代码”子文件夹和“实验报告”子文件