

# 一 实验名称：朴素贝叶斯

前面实验过程中我们使用的分类器都是直接作出分类决策，给出该数据实例属于哪一类这类问题的明确答案。不过分类器有时会产生错误结果，这是要求分类器给出一个最优的类别猜测结果，同时给出这个猜测的概率估计值。

# 二 数据描述

Lenses 数据集包含很多患者眼部状况的观察条件以及医生推荐的隐形眼镜类型。Lenses.txt 数据集一共 24 条数据，每条数据为一个患者的观察情况（age, prescript, astigmatic, tearRate）。隐形眼镜类型包括硬材质、软材质以及不适合佩戴隐形眼镜(hard, soft, no lenses)。

# 三 实验内容

- 1. 数据探索性分析  
利用第一次实验中的因子分析方式，探索性分析数据特点和分布。
- 2. 分类  
利用朴素贝叶斯进行分类。
- 3. 利用准确率、召回率和 F1 值衡量模型

# 四 实验环境和工具

系统环境：windows all/linux  
程序语言：Python  
编译工具：Pycharm  
第三方包：numpy, matplotlib, seaborn, sklearn

# 五 相关知识

表 2.1 分类结果混淆矩阵

真实情况	预测结果	
	正例	反例
正例	TP (真正例)	FN (假反例)
反例	FP (假正例)	TN (真反例)

### 1. 准确率

$$P = \frac{TP}{TP + FP}$$

### 2. 召回率

$$P = \frac{TP}{TP + FN}$$

### 3. F1 值

$$F1 = \frac{2 \times P \times R}{P + R}$$