1. **贝叶斯分类器作业题**

**1. 单特征分类器实验 The Chosen Four、PLD、呵，recognition、NSAS、追风**

1. 利用高斯公式，分别生成鲈鱼（1类）和鲑鱼（2类）长度的10000个数值
2. 利用每个类的5000个数值，求出概率密度曲线（可以利用直方图）
3. 假设先验概率相同，均为0.5，计算另外5000个数据的分类结果
4. 计算你的分类结果的正确率和错误率

（包括1类分为2类的错误率和2类分为1类的错误率）

1. 评估你的错误率和理论推导的错误率的异同

**2. 狼来了**

**呵，pattern、老师很可爱、2018ML666、Peppa Pig、机器学习业余专家**

一些合理的假设: 事件A表示:“小孩说谎”；事件B表示:“小孩可信”。

村民起初对这个小孩的信任度为0.8,即 P(B)=0.8。

我们认为可信的孩子说谎的可能性为0.1。即 P(A/B)=0.1。

不可信的孩子说谎的可能性为0.5,即 P(A/^B)=0.5 (用 ^B 表示B的对立事件)。

第一次说谎的后果: 第一次村民上山打狼发现狼没有来：即小孩说了谎(A)，根据贝叶斯论，这时村民觉得小孩的可信程度变为：

P(B/A)= [P(A/B)\*P(B)]/[P(A/B)\*P(B)+P(A/^B)\*P(^B)]

= (0.1\*0.8)/[0.1\*0.8+0.5\*(1-0.8)]

= 0.444

这表明村民上了一次当以后对小孩的信任程度由原来的0.8降到0.444了，即

P(B)=0.444, P(^B)=1- P(B)=0.556.

第二次说谎的后果:

第二次村民上山打狼发现狼没有来这时村民觉得小孩的可信程度变为：

P(B/A)= [P(A/B)\*P(B)] / [P(A/B)\*P(B)+P(A/^B)\*P(^B)]

= (0.1\*0.444)/[0.1\*0.444+0.5\*(1-0.444]

= 0.138

这表明村民经过两次上当，对这个小孩的信任程度已经从0.8降到了0.138。如此低的可信度，别说狼来了，喊华南虎来了都不管用了。

那么，如果这个孩子要改邪归正...

需要多少次才能把信任度提高到80%？

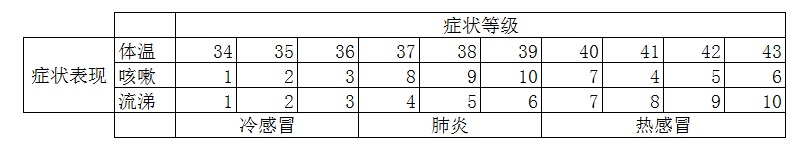
**3. 请理解CORK\_STOPPERS.XLS数据的基本内容，假设每个特征都服从单峰正态分布，计算每个特征的均值和方差，然后对根据每类特征对样本进行分类(假设只有1，2两类)，求出准确率。训练样本和测试样本可以一样。哪个特征的分类性能最好？ Inception、666、818、XBM、007、chili**

**4. 朴素贝叶斯分类器诊断病情**

**奇妙能力组、Alpha 3、777、哈库拉玛塔塔、组名**

条件A：一个人感冒了；条件B：这个人会发烧

朴素贝叶斯根据这10组数据进行训练，然后输入一位病人的数据，自动进行诊断，以下是训练数据：



然后输入一位病人的数据，自动进行诊断，并对过程及结果进行可视化。