计算机与信息工程学院实验报告

••••••••••••••••••••••••••••••••• 密 ••••••••••••••••••••••••••••••••• 封 ••••••••••••••••••••••••••••••••• 线 •••••••••••••••••••••••••••••••••

姓名： 学号： 201 专业：计算机科学与技术

年级： 2020 课程： 人工智能导论

主讲教师： 刘扬 辅导教师： 刘扬

时间：2021年 10 月 22 日下午17时至18时，地点：基础实验中心网络实训室

实验题目： 统计学习实验

实验目的： 本实验采用概率和统计方法，构建统计模型，从数据中挖掘到感兴趣的信息，实现智 能预测的机制，达到课程目标的要求。

实验环境（硬件和软件） 64位i5-1035G1 CPU Windows10中文专业版 编译器：DEVC++

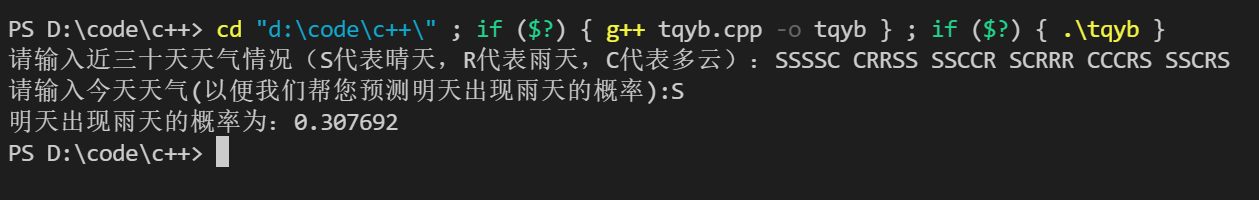
实验内容：

假设晴天(Sunny)、阴天(Cloudy)、下雨(Rainy)分别用 S,C,R 表示，过去 30 天天气情况 分别为 SSSSC CRRSS SSCCR SCRRR CCCRS SSCRS，选择合理的预测模型，根据历史数 据，用统计学习方法，预测某地如果今天分别为 S,C,R，第二天下雨的概率。

实验步骤：

1. 创建工程
2. 根据朴素贝叶斯公式，编写函数计算预测概率。
3. 测试多组样例，输入最近30天的天气，以及今天的天气，预测明天天气。

实验数据记录：

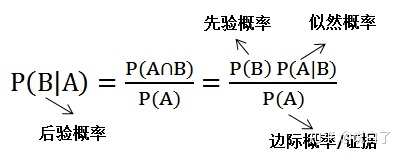


问题讨论：

1. 如何初步确定事件的发生？ 事件=确定事件+随机事件  
   确定事件=必然事件+不可能事件
2. 为了进一步得到天气预测的概率，准确预测天气发生的概率，应该选择哪一个公式来计算？

利用贝叶斯公式对天气的发生概率进行推测：

得到了条件概率的一般定义：P(A|B)=P(AB)/P(B)。



1. 关于原始频率的计算：

分别统计：昨天的具体天气，明天的具体天气的次数，再分别除以总天气数，得出各种情况的频率。

1. 关于贝叶斯函数的实现细节：

采用c++编写，要处理前一天是雨天，后一天是某天的出现频数。我们可以置一个flag，动态更新这个flag，当这一天是雨天时，将这个flag置1，如此可以识别这种情况。

另外，需要注意int除法和浮点除法的使用，在计算概率时候一定要使用浮点除法。

1. 处理数据集合的方法：  
   根据前30天的天气情况，编写一个字符串，用来预测概率