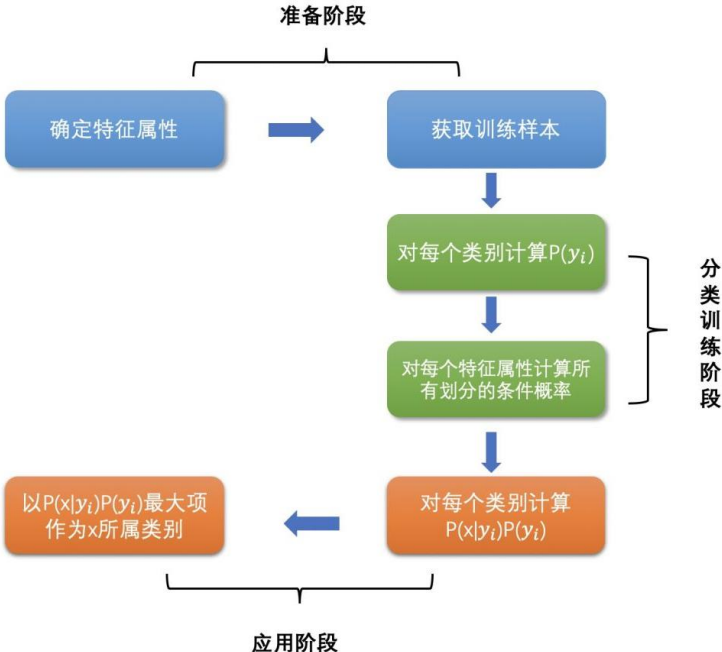


1	<p>写出全概率公式&贝叶斯公式</p> $p(A B) = \frac{P(A \cap B)}{P(B)} \Rightarrow p(A \cap B) = p(A B) \times p(B)$ $\Rightarrow p(A \cap B) = P(B A) \times P(A)$ <p>因此条件概率公式:</p> $p(A B) = \frac{P(A \cap B)}{P(B)} = \frac{p(B A) \times P(A)}{p(B)}$ <p>假如事件 A 与 B 相互独立, 那么:</p> $p(A \cap B) = P(A) \times P(B)$ <p>设事件 L_1, L_2, \dots 是一个完备事件组, 则对于任意一个事件 C, 则有全概率公式:</p> $p(C) = p(L_1)p(C L_1) \cdots p(L_n)p(C L_n) = \sum_{i=1}^n p(L_i)p(C L_i)$ <p>在已知条件概率和全概率的基础上, 贝叶斯公式如下:</p> $p(L_k C) = \frac{p(C L_k) \times p(L_k)}{p(C)}$ $\Rightarrow p(L_k C) = \frac{p(C L_k) \times p(L_k)}{\sum_{i=1}^n p(L_i) \times p(C L_i)}$	从形式上理解两个公式之间的关系
2	<p>朴素贝叶斯为什么“朴素 naive”?</p> <p>朴素贝叶斯模型, 全称为: Naive Bayesian Model。它是简单的, 缺乏 worldly experience 的。朴素贝叶斯模型假设样本特征彼此独立, 没有相关关系。正如我们所知, 这个假设在现实世界中是很不真实的, 因此, 说朴素贝叶斯真的很“朴素”。这个假设现实中基本上不存在, 但特征相关性很小的实际问题还是很多的, 所以这个模型仍然能够工作得很好</p>	朴素一词其实是个形容词, 在这里可同理想化等同
3	<p>朴素贝叶斯有没有超参数可以调?</p> <p>朴素贝叶斯是没有超参数可以调的, 朴素贝叶斯是根据训练集进行分类, 分类出来的结果基本上就是确定了, 拉普拉斯估计器不是朴素贝叶斯中的参数, 不能通过拉普拉斯估计器来对朴素贝叶斯调参。</p>	根据公式来判断

4	<p>朴素贝叶斯的工作流程是怎样的？</p> <p>朴素贝叶斯的工作流程可以分为三个阶段进行，分别是准备阶段、分类器训练阶段和应用阶段。</p> <p>1. 准备阶段：这个阶段的任务是为朴素贝叶斯分类做必要的准备，主要工作是根据具体情况确定特征属性，并对每个特征属性进行适当划分，去除高度相关性的属性（如果两个属性具有高度相关性的话，那么该属性将会在模型中发挥了 2 次作用，会使得朴素贝叶斯所预测的结果向该属性所希望的方向偏离，导致分类出现偏差），然后由人工对一部分待分类项进行分类，形成训练样本集合。这一阶段的输入是所有待分类数据，输出是特征属性和训练样本。（这一阶段是整个朴素贝叶斯分类中唯一需要人工完成的阶段，其质量对整个过程将有重要影响。）</p> <p>2. 分类器训练阶段：这个阶段的任务就是生成分类器，主要工作是计算每个类别在训练样本中的出现频率及每个特征属性划分对每个类别的条件概率估计，并将结果记录。其输入是特征属性和训练样本，输出是分类器。这一阶段是机械性阶段，根据前面讨论的公式可以由程序自动计算完成。</p> <p>3. 应用阶段：这个阶段的任务是使用分类器对待分类项进行分类，其输入是分类器和待分类项，输出是待分类项与类别的映射关系。这一阶段也是机械性阶段，由程序完成。</p>  <pre> graph TD subgraph 准备阶段 A[确定特征属性] --> B[获取训练样本] end B --> C[对每个类别计算P(y_i)] C --> D[对每个特征属性计算所有划分的条件概率] D --> E[对每个类别计算P(x y_i)P(y_i)] E --> F[以P(x y_i)P(y_i)最大项作为x所属类别] subgraph 应用阶段 F end style C fill:#90EE90 style D fill:#90EE90 style E fill:#FFDAB9 style F fill:#FFDAB9 </pre>	<p>从数据准备, 训练和应用三个阶段来理解</p>
5	<p>朴素贝叶斯对异常值敏不敏感？</p> <p>朴素贝叶斯对异常值不敏感。所以在进行数据处理时，我们可以不去除异常值，因为保留异常值可以保持朴素贝叶斯算法的整体精度，而去除异常值则可能在预测的过程中由于失去部分异常值导致模型的泛化能力下降。</p>	<p>从朴素贝叶斯的工作流程来理解</p>

