Eric Zhang's Tech Blog (<http://leo2sk.cnblogs.com/>)

带你理解朴素贝叶斯分类算法



忆臻

哈尔滨工业大学 计算机科学与技术博士在读

关注他

2,410 人赞同了该文章

贝叶斯分类是一类分类算法的总称，这类算法均以贝叶斯定理为基础，故统称为贝叶斯分类。而朴素贝叶斯分类是贝叶斯分类中最简单，也是常见的一种分类方法。这篇文章我尽可能用直白的话语总结一下我们学习会上讲到的朴素贝叶斯分类算法，希望有利于他人理解。

分类问题综述

对于分类问题，其实谁都不会陌生，日常生活中我们每天都进行着分类过程。例如，当你看到一个人，你的脑子下意识判断他是学生还是社会上的人；你可能经常会走在路上对身旁的朋友说“这个

人一看就很有钱、”之类的话，其实这就是一种分类操作。

既然是贝叶斯分类算法，那么**分类的数学描述**又是什么呢？

从数学角度来说，分类问题可做如下定义：已知集合 $C = y_1, y_2, \dots, y_n$ 和 $I = x_1, x_2, x_3, \dots, x_n$ ，确定映射规则 $y = f()$ ，使得任意 $x_i \in I$ 有且仅有一个 $y_i \in C$ ，使得 $y_i \in f(x_i)$ 成立。

其中C叫做类别集合，其中每一个元素是一个类别，而I叫做项集合（特征集合），其中每一个元素是一个待分类项，f叫做分类器。**分类算法的任务就是构造分类器f。**

分类算法的内容是要求给定特征，让我们得出类别，这也是所有分类问题的关键。那么如何由指定特征，得到我们最终的类别，也是我们下面要讲的，每一个不同的分类算法，对应着不同的核心思想。

本篇文章，我会用一个具体实例，对朴素贝叶斯算法几乎所有的重要知识点进行讲解。

朴素贝叶斯分类

那么既然是朴素贝叶斯**分类算法**，它的核心算法又是什么呢？

是下面这个贝叶斯公式：

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

换个表达形式就会明朗很多，如下：

$$p(\text{类别}|\text{特征}) = \frac{p(\text{特征}|\text{类别})p(\text{类别})}{p(\text{特征})}$$

我们最终求的p(类别|特征)即可！就相当于完成了我们的任务。

例题分析

下面我先给出例子问题。

给定数据如下：

帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁



现在给我们的问题是，如果一对男女朋友，男生想女生求婚，男生的四个特点分别是不帅，性格不好，身高矮，不上进，请你判断一下女生是嫁还是不嫁？

这是一个典型的分类问题，转为数学问题就是比较 $p(\text{嫁} | (\text{不帅、性格不好、身高矮、不上进}))$ 与 $p(\text{不嫁} | (\text{不帅、性格不好、身高矮、不上进}))$ 的概率，谁的概率大，我就能给出嫁或者不嫁的答案！

这里我们联系到朴素贝叶斯公式：

$$p(\text{嫁} | \text{不帅、性格不好、身高矮、不上进}) = \frac{p(\text{不帅、性格不好、身高矮、不上进} | \text{嫁}) * p(\text{嫁})}{p(\text{不帅、性格不好、身高矮、不上进})}$$

我们需要求 $p(\text{嫁} | (\text{不帅、性格不好、身高矮、不上进}))$ ，这是我们不知道的，但是通过朴素贝叶斯公式可以转化为好求的三个量， $p(\text{不帅、性格不好、身高矮、不上进} | \text{嫁})$ 、 $p(\text{不帅、性格不好、身高矮、不上进})$ 、 $p(\text{嫁})$ （至于为什么能求，后面会讲，那么就太好了，将待求的量转化为其它可求的值，这就相当于解决了我们的问题！）

朴素贝叶斯算法的朴素一词解释

那么这三个量是如何求得？

是根据已知训练数据统计得来，下面详细给出该例子的求解过程。

回忆一下我们要求的公式如下：

$$p(\text{嫁} | \text{不帅、性格不好、身高矮、不上进}) = \frac{p(\text{不帅、性格不好、身高矮、不上进} | \text{嫁}) * p(\text{嫁})}{p(\text{不帅、性格不好、身高矮、不上进})}$$

那么我只要求得 $p(\text{不帅、性格不好、身高矮、不上进} | \text{嫁})$ 、 $p(\text{不帅、性格不好、身高矮、不上进})$ 、 $p(\text{嫁})$ 即可，好的，下面我分别求出这几个概率，最后一比，就得到最终结果。

$p(\text{不帅、性格不好、身高矮、不上进} | \text{嫁}) = p(\text{不帅} | \text{嫁}) * p(\text{性格不好} | \text{嫁}) * p(\text{身高矮} | \text{嫁}) * p(\text{不上进} | \text{嫁})$ ，那么我就要分别统计后面几个概率，也就得到了左边的概率！

等等，为什么这个成立呢？学过概率论的同学可能有感觉了，这个等式成立的条件需要特征之间相互独立吧！

对的！这也就是为什么朴素贝叶斯分类有朴素一词的来源，朴素贝叶斯算法是假设各个特征之间相互独立，那么这个等式就成立了！

但是为什么需要假设特征之间相互独立呢？

1、我们这么想，假如没有这个假设，那么我们对右边这些概率的估计其实是不可做的，这么说，我们这个例子有4个特征，其中帅包括{帅，不帅}，性格包括{不好，好，爆好}，身高包括{高，矮，中}，上进包括{不上进，上进}，那么四个特征的联合概率分布总共是4维空间，总个数为 $2 * 3 * 3 * 2 = 36$ 个。

24个，计算机扫描统计还可以，但是现实生活中，往往有非常多的特征，每一个特征的取值也是非常之多，那么通过统计来估计后面概率的值，变得几乎不可做，这也是为什么需要假设特征之间独立的原因。

2、假如我们没有假设特征之间相互独立，那么我们统计的时候，就需要在整个特征空间中去寻找，比如统计 $p(\text{不帅、性格不好、身高矮、不上进} | \text{嫁})$ ，

我们就需要在嫁的条件下，去找四种特征全满足分别是不帅，性格不好，身高矮，不上进的人的个数，这样的话，由于数据的稀疏性，很容易统计到0的情况。这样是不合适的。

根据上面两个原因，朴素贝叶斯法对条件概率分布做了条件独立性的假设，由于这是一个较强的假设，朴素贝叶斯也由此得名！这一假设使得朴素贝叶斯法变得简单，但有时会牺牲一定的分类准确率。

好的，上面我解释了为什么可以拆成分开连乘形式。那么下面我们就开始求解！



我们将上面公式整理一下如下：

$$p(\text{嫁}|\text{不帅、性格不好、身高矮、不上进}) = \frac{p(\text{不帅、性格不好、身高矮、不上进}|\text{嫁}) * p(\text{嫁})}{p(\text{不帅、性格不好、身高矮、不上进})}$$
$$= \frac{p(\text{不帅}|\text{嫁}) * p(\text{性格不好}|\text{嫁}) * p(\text{身高矮}|\text{嫁}) * p(\text{不上进}|\text{嫁}) * p(\text{嫁})}{p(\text{不帅}) * p(\text{性格不好}) * p(\text{身高矮}) * p(\text{不上进})}$$

下面我将一个一个的进行统计计算（在数据量很大的时候，根据中心极限定理，频率是等于概率的，这里只是一个例子，所以我就进行统计即可）。

$p(\text{嫁}) = ?$

首先我们整理训练数据中，嫁的样本数如下：

帅？	性格好？	身高？	上进？	嫁与否
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁

则 $p(\text{嫁}) = 6/12$ （总样本数） = $1/2$

$p(\text{不帅}|\text{嫁}) = ?$ 统计满足样本数如下：

帅？	性格好？	身高？	上进？	嫁与否
不帅	好	高	上进	嫁
不帅	好	中	上进	嫁
不帅	不好	高	上进	嫁

则 $p(\text{不帅}|\text{嫁}) = 3/6 = 1/2$

$p(\text{性格不好}|\text{嫁}) = ?$ 统计满足样本数如下：

帅？	性格好？	身高？	上进？	嫁与否
不帅	不好	高	上进	嫁

则 $p(\text{性格不好}|\text{嫁}) = 1/6$

$p(\text{矮}|\text{嫁}) = ?$ 统计满足样本数如下：

帅？	性格好？	身高？	上进？	嫁与否
帅	好	矮	上进	嫁

则 $p(\text{矮}|\text{嫁}) = 1/6$

$p(\text{不上进}|\text{嫁}) = ?$ 统计满足样本数如下：

帅？	性格好？	身高？	上进？	嫁与否
帅	好	高	不上进	嫁

则 $p(\text{不上进}|\text{嫁}) = 1/6$



下面开始求分母，p(不帅)，p(性格不好)，p(矮)，p(不上进)

统计样本如下：

帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁

不帅统计如上红色所示，占4个，那么p(不帅) = 4/12 = 1/3

帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁

性格不好统计如上红色所示，占4个，那么p(性格不好) = 4/12 = 1/3

帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁

身高矮统计如上红色所示，占7个，那么p(身高矮) = 7/12

帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁

不上进统计如上红色所示，占4个，那么 $p(\text{不上进}) = 4/12 = 1/3$

到这里，要求 $p(\text{不帅、性格不好、身高矮、不上进|嫁})$ 的所需项全部求出来了，下面我带入进去即可，

$$\begin{aligned}
 p(\text{嫁}|\text{不帅、性格不好、身高矮、不上进}) &= \frac{p(\text{不帅、性格不好、身高矮、不上进|嫁}) * p(\text{嫁})}{p(\text{不帅、性格不好、身高矮、不上进})} \\
 &= \frac{p(\text{不帅|嫁}) * p(\text{性格不好|嫁}) * p(\text{身高矮|嫁}) * p(\text{不上进|嫁}) * p(\text{嫁})}{p(\text{不帅}) * p(\text{性格不好}) * p(\text{身高矮}) * p(\text{不上进})} \\
 &= (1/2 * 1/6 * 1/6 * 1/6 * 1/2) / (1/3 * 1/3 * 7/12 * 1/3)
 \end{aligned}$$

下面我们根据同样的方法来求 $p(\text{不嫁|不帅，性格不好，身高矮，不上进})$ ，完全一样的做法，为了方便理解，我这里也走一遍帮助理解。首先公式如下：

$$\begin{aligned}
 p(\text{不嫁}|\text{不帅、性格不好、身高矮、不上进}) &= \frac{p(\text{不帅、性格不好、身高矮、不上进|不嫁}) * p(\text{不嫁})}{p(\text{不帅、性格不好、身高矮、不上进})} \\
 &= \frac{p(\text{不帅|不嫁}) * p(\text{性格不好|不嫁}) * p(\text{身高矮|不嫁}) * p(\text{不上进|不嫁}) * p(\text{不嫁})}{p(\text{不帅}) * p(\text{性格不好}) * p(\text{身高矮}) * p(\text{不上进})}
 \end{aligned}$$

下面我也一个一个来进行统计计算，这里与上面公式中，分母是一样的，于是我们分母不需要重新统计计算！

$p(\text{不嫁}) = ?$ 根据统计计算如下（红色为满足条件）：

帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁

则 $p(\text{不嫁}) = 6/12 = 1/2$

$p(\text{不帅|不嫁}) = ?$ 统计满足条件的样本如下（红色为满足条件）：



帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁

则 $p(\text{不帅}|\text{不嫁}) = 1/6$

$p(\text{性格不好}|\text{不嫁}) = ?$ 据统计计算如下（红色为满足条件）：

帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁

则 $p(\text{性格不好}|\text{不嫁}) = 3/6 = 1/2$

$p(\text{矮}|\text{不嫁}) = ?$ 据统计计算如下（红色为满足条件）：

帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁

则 $p(\text{矮}|\text{不嫁}) = 6/6 = 1$

$p(\text{不上进}|\text{不嫁}) = ?$ 据统计计算如下（红色为满足条件）：



帅？	性格好？	身高？	上进？	嫁与否
帅	不好	矮	不上进	不嫁
不帅	好	矮	上进	不嫁
帅	好	矮	上进	嫁
不帅	好	高	上进	嫁
帅	不好	矮	上进	不嫁
帅	不好	矮	上进	不嫁
帅	好	高	不上进	嫁
不帅	好	中	上进	嫁
帅	好	中	上进	嫁
不帅	不好	高	上进	嫁
帅	好	矮	不上进	不嫁
帅	好	矮	不上进	不嫁

则 $p(\text{不上进}|\text{不嫁}) = 3/6 = 1/2$

那么根据公式：

$$p(\text{不嫁}|\text{不帅、性格不好、身高矮、不上进}) = \frac{p(\text{不帅、性格不好、身高矮、不上进}|\text{不嫁}) * p(\text{不嫁})}{p(\text{不帅、性格不好、身高矮、不上进})}$$
$$= \frac{p(\text{不帅}|\text{不嫁}) * p(\text{性格不好}|\text{不嫁}) * p(\text{身高矮}|\text{不嫁}) * p(\text{不上进}|\text{不嫁}) * p(\text{不嫁})}{p(\text{不帅}) * p(\text{性格不好}) * p(\text{身高矮}) * p(\text{不上进})}$$

$$p(\text{不嫁}|\text{不帅、性格不好、身高矮、不上进}) = ((1/6 * 1/2 * 1 * 1/2) * 1/2) / (1/3 * 1/3 * 7/12 * 1/3)$$

很显然 $(1/6 * 1/2 * 1 * 1/2) > (1/2 * 1/6 * 1/6 * 1/6 * 1/2)$

于是有 $p(\text{不嫁}|\text{不帅、性格不好、身高矮、不上进}) > p(\text{嫁}|\text{不帅、性格不好、身高矮、不上进})$

所以我们根据朴素贝叶斯算法可以给这个女生答案，是不嫁！！！！

朴素贝叶斯分类的优缺点

优点：

- (1) 算法逻辑简单,易于实现
- (2) 分类过程中时空开销小

缺点：

理论上，朴素贝叶斯模型与其他分类方法相比具有最小的误差率。但是实际上并非总是如此，这是因为朴素贝叶斯模型假设属性之间相互独立，这个假设在实际应用中往往是不成立的，在属性个数比较多或者属性之间相关性较大时，分类效果不好。

而在属性相关性较小时，朴素贝叶斯性能最为良好。对于这一点，有半朴素贝叶斯之类的算法通过考虑部分关联性适度改进。

整个例子详细的讲解了朴素贝叶斯算法的分类过程，希望对大家的理解有帮助~

参考：李航博士《统计学习方法》

算法杂货铺--分类算法之朴素贝叶斯分类(Naive Bayesian classification)



致谢：德川，皓宇，继豪，施琦

编辑于 2017-04-11 00:00

「真诚赞赏，手留余香」

赞赏

10 人已赞赏



[机器学习](#) [深度学习 \(Deep Learning\)](#) [自然语言处理](#)

文章被以下专栏收录



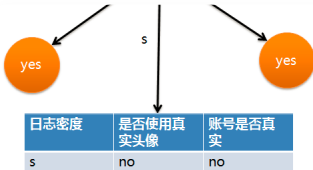
机器学习算法与自然语言处理
公号[机器学习算法与自然语言处理] 微信号yizhennotes

推荐阅读



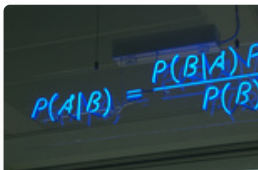
朴素贝叶斯分类实例-单词纠正问题

忆臻 发表于机器学习算...



理解朴素贝叶斯分类的拉普拉斯平滑

忆臻



朴素贝叶斯最大似然估计资料)

贪心学院

132 条评论

切换为时间排序

写下你的评论...



之虫备紫冥

2018-04-26

首先写的真的很好，然后就是措辞上可能有点问题，条件独立和事件独立是不一样的，条件独立是conditional independence，和特征之间的independent 不一样，而且conditional independence应该是一种假设，应该无法推导出来。应该写成特征|类别的条件独立（conditional independence），不能写成特征之间独立（independent）。这是我的想法。

24



SwarmLover 回复 之虫备紫冥

2021-03-05

对的。独立可以推出条件独立，但是条件独立不一定独立。因此条件独立这个条件是更弱的，更加合理。

1



中科大哈工第一 回复 SwarmLover

2021-04-22

独立不能推出条件独立

👍 赞



展开其他 1 条回复



励志橙

2019-10-20

将分母改成P (嫁) P (不帅, 不好, 矮, 不上进|嫁) + P (不嫁) P (不帅, 不好, 矮, 不上进|不嫁) 后计算结果: P (嫁|不帅, 不好, 矮, 不上进) = 1/19, P (不嫁|不帅, 不好, 矮, 不上进) = 18/19。

谢谢你, 数学不好的我都能看懂! 你的文章都写得好好! 作者你超级帅



👍 15



Xing Zhan 回复 励志橙

2020-07-25

“不帅, 不好, 矮, 不上进” 这这情况在表格里都不存在 (即样本数量为0), 你怎么统计计算 “P (不帅, 不好, 矮, 不上进|嫁)” 和 “P (不帅, 不好, 矮, 不上进|不嫁)” 的数值呢?

👍 赞



攀条折其荣 回复 Xing Zhan

2021-01-01

用拉普拉斯平滑 (贝叶斯估计) 吧

👍 2

展开其他 1 条回复



风云劫

2017-04-11

拉普拉斯平滑没讲到?

👍 10



忆臻 (作者) 回复 风云劫

2017-04-11

我打算下篇文章再用这个例子讲~

👍 2



Eric 回复 风云劫

2017-04-11

好巧, 竟然在这儿遇见你

👍 1



Anish Hui

2020-04-10

文章写得好棒👍, 不过频率等于概率是大数定理吧

👍 5



本人依然没救

2019-10-08

讲的很棒 可是例子看着看着眼睛里进了砖头

👍 5



九日九日 回复 本人依然没救

02-23

Hhhhhhha, 我笑出声

👍 赞



凝舞

2017-12-26

讲的很朴素, 全都看懂了, 谢谢!

👍 5



EliC

2017-06-22

在不帅、性格不好、身高矮、不上进的条件下, 嫁与不嫁是两个对立的事件, 为什么 $p(\text{不嫁}|\text{不帅、性格不好、身高矮、不上进}) + p(\text{嫁}|\text{不帅、性格不好、身高矮、不上进}) \neq 1$?

👍 3



面壁者高首 回复 EliC

2019-01-10

是不是因为概率在数据量大的时候才可以近似等于频率，这个导致的

👍 1



无痕 回复 EliC

2019-03-06

应该是答主分母计算错了，特征条件独立并不是特征独立，计算分母的联合特征概率时应该用贝叶斯全概率公式，这样就可转化为特征条件独立进行计算

$p(\text{不帅、性格不好、不上进, 矮}) = p(\text{嫁}) * p(\text{不帅、性格不好、不上进, 矮}|\text{嫁}) + p(\text{不嫁}) * p(\text{不帅、性格不好、不上进, 矮}|\text{不嫁})$

👍 39

展开其他 3 条回复



孙勇

2019-05-17

是根据概率论的大数定律，不是根据中心极限定理吧

👍 9



aboutlatte 回复 孙勇

2020-06-22

我赞同你的看法，中心极限定理好像是对于抽样的均值而言的。

👍 1



武理杜兰特 回复 孙勇

2021-04-19

确实是根据大数定律，我还特意翻了一下书

👍 赞



知 一百八十天不改名

2018-05-16

很好的解释了“**为什么需要假设特征之间相互独立**”，赞一个！

👍 9



Leechard

2018-02-16

分母部分的计算是有问题的。特征之间的独立与特征之间的条件独立是两码事。朴素贝叶斯假设是特征之间条件独立，分子计算是对的，分母的计算默认了特征独立，我认为有误。

👍 7



忆臻 (作者) 回复 Leechard

2018-02-16

你好，新年快乐,具体是哪块有问题？

👍 赞



Leechard 回复 忆臻 (作者)

2018-02-16

新年好！例子的分母计算。 $p(\text{不帅、性格不好、身高矮、不上进}) = p(\text{不帅}) * p(\text{性格不好}) * p(\text{身高矮}) * p(\text{不上进})$ ，这一步的推导。

👍 3

展开其他 3 条回复



三一

2017-04-11

非常漂亮的讲述，物质和精神都赞一个！

👍 3



忆臻 (作者) 回复 三一

2017-04-11

很开心对你有帮助，谢谢~

👍 1



何晔

2017-04-11

如果某一分类没有数据怎么办？那他的概率就是零了。

👍 3



忆臻 (作者) 回复 何晔

2017-04-11

你好，下篇文章我会讲平滑~

👍 3





Jingyang Zhang

2021-02-09

呜呜呜讲的真好！终于明白了

👍 2



黑暗中的灯光

2020-08-21

是你写的如此通俗易懂，还是李航博士《统计学习方法》写的如此通俗易懂？比上来就写公式的强了好多。

👍 2



小贝壳

2019-04-21

就喜欢这种能够以通俗易懂地方式讲问题的作者，很棒

👍 2



大风哥

2021-11-10

讲的真好

👍 1



濼破天

2020-08-24

学习了，谢谢

👍 1



A俊俊

2020-06-30

讲的很好！赞一个

👍 1



Chilan Yuk 🏆

2020-02-01

喜欢作者的写作风格，通俗易懂，简单有趣

👍 1



Ahern

2019-12-17

棒

👍 1