贝叶斯算法

贝叶斯要解决的问题：

正向概率：已知黑盒中有a个红球，b个黄球，那么拿出一个球是黄球的概率多大？这种问题就是正向概率，由已知的条件俩推测可能出现的结果。

逆向概率：同样是一个黑盒，但不清楚红球与黄球的比例，实验10次，每次拿出一个球后又放回，通过这10次实验结果，来推测黑盒中的红球黄球比例。这种问题叫逆向概率，贝叶斯算法要解决的就是这种问题，这也是我们日常中更容易遇到的问题，通常要从结果来推测造成结果的原因。即从看到的现象中推测黑盒中的情况

举例：

男生与女生的比例为60%和40%，且男生只穿长裤，女生一半穿长裤，一半穿裙子。

男生，女生总数为W个，则穿长裤的男生的人数：N1 = W \* P(boy) \*P(pants/boy)

P(boy) = 60 %（学生中男生的概率）

P(Pants/boy) 条件概率，在学生是男生的条件下，穿长裤的概率是100%

女生：

N2 = W \* P(boy) \*P(pants/boy)

P(girl) = 40 %

P(pants/girl) = 50 %

逆向概率：走来一个穿长裤的学生，判断性别

即在穿长裤的条件下，求出学生是女生或男生的概率。

1. 求出穿长裤的男生人数 N1
2. 求出穿长裤的女生人数 N2
3. 学生是男生的概率为p(boy/pants) = N1/(N1+N2) = P(pants/boy) / P(pants)
4. 学生是女生的概率为P(girl/pants) = N2 / (N1+N2) = P(pants/girl)/P(pants)

通常逆向概率的场景是我们发现了B这个条件，要通过B这个条件来推测A的概率。但通常情况下P(A/B)这个概率是不太好求的，但是P(B/A)这个概率很好求，即知道条件A，推测B的概率。

在上例中对应的就是我们知道学生总人数，知道男女生比例，同时也知道男女生中穿长裤的比例，因此，可以从知晓这个学生的性别这个A条件中，很容易地可以算出基于学生性别的条件A的B事件，即推测这个学生穿长裤的概率P(B/A)。

但如果给出这个学生是穿长裤这个B条件，就比较难于从现有的信息里推断出学生的性别这个A条件。这就是逆向概率的应用场景，通过贝叶斯公式可以将一个比较难求的概率P(A/B)转换为了很好求的概率P(B/A)的式子

所以推导出逆向概率的公式如下：

P(A/B) = P(B/A)P(A) / P(B)

实例：

拼写纠正的实例，场景：用户输入了不在词典中的单词，我们需要根据输入猜测用户的想法。 即在用户输入的单词这个条件B上，我们来推测可能的单词列表A，然后比较概率，概率最大的即是最有可能的单词。

因此，通过贝叶斯公式，我们想到将P(A/B)转换为P(B/A)的式子，因为在可能的单词列表A这个条件上，我们可以通过键盘上键位的距离计算出将可能的单词列表A中每个单词输错为B的概率，即P(B/A),由此可以转化为P(B/A)的式子。

P(A/B) = P(B/A) \* P(A) / P(B)

P(A)是每个可能的单词在词典中的概率，即先验概率，因为一般都是已知的。P(B)是用户输入某个单词的概率，是随机的，而且在后面概率的比较中是可以约去的，对于结果没有影响。

对于单词了列表A中的不同单词A1，A2…，P(B)都是一样的，因此被约分掉了。所以不同Ai的比较中，只有分子是有影响的。即

P(A/B) 正比于 P(A) \* P(B/A)

理解为：一个猜测的好坏与否，跟这个猜测本身的独立概率（即这个猜测的单词在你曾经输入的词库中的占比）和这个猜测转化为观测数据的可能性成正比。