人工智能基础作业 9

张磊 2017K8009922027

13.15 解：由题可知，

P(阳性|患病) = 0.99；

P(阴性|不患病) = 0.99；

P(患病) = 0.0001；

而我关心的是 P(患病|阳性)的值，即我确实患有这种病的概率为；

$$P(患病|阳性) = \frac{P(阳性|患病) * P(患病)}{P(阳性|患病) * P(患病) + P(阳性|不患病) * P(不患病)}$$
$$= \frac{0.99 * 0.0001}{0.99 * \ 0.0001 + 0.01 * 0.9999}$$
$$= 0.009804$$

可见，只要 P(患病)足够小，即“这种病很罕见”，我患病的概率就很小；

13.21 解：设 B 事件为出租车是蓝色，A 事件为出租车看起来是蓝色；

a.

那么，P(A|B) = 0.75，P(~A|~B) = 0.75；

我们关心的是，当看到出租车为蓝色的条件下，出租车确实为蓝色或者实际是绿色的概率；

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

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

P(B|A) ∝ P(A|B) * P(B) ∝ 0.75 * P(B)；

P(~B|A) ∝ P(A | ~B)*P(~B) ∝ 0.25 * (1 – P(B))；

所以，只有当 P(B)超过 0.25 的情况下，我们才有可能得出肇事车辆最可能的颜色，目前由于无法确认 P(B)的值，所以不能判断；

b.

如果 10 辆出租车种有 9 辆是绿色的，那么 P(B) = 0.1 < 0.25，所以，肇事车辆更有可能是绿色的；

13.22 解：

a. 该模型包含先验概率 P(Category)，和条件概率 P(Wordi|Category)；任给定一个类别 c，P(Category = c)表示属于该类的文件占所有文件的比例；而，P(Wordi = True|Category = c)表示 c 类文件中含有 Worldi 的比例；

b. 通过提取新文档中的词语，计算新文档中的关键词汇所占的比例，将其与已经分类的文档进行比较，选择比例相近的类别进行归类；

c. 不合理，因为某些特定的词组在任何给定的文档中出现的频率都比将它拆开后，几个词在文档中出现的概率相乘评估的频率要高；比如“人工智能”出现的频率就比“人工”和“智能”出现的概率相乘评估的频率还要高；

13.23 解:

a. 变量 $P_{i,j}$ 和 $P_{k,l}$ 不再独立，通常:

$P(P_{i,j} = true | P_{k,l} = true) < P(P_{i,j} = true | P_{k,l} = false)$；

b. 联合分布概率为，给所有恰好包含 3 个陷阱的赋值以相同的概率，给其他赋值以以 0 的概率；在 15 个方格中包含 3 个陷阱，所以，恰好包含 3 个陷阱的赋值概率为 $1/C_{15,3} = 1/455$;

c. 由中文版教材图 13.6 (a)可得;

[1,3]有陷阱可分为 3 种情况，第 1 种，3 个陷阱在[1,3],[2,2],[3,1]中，第 2 种，陷阱在[1,3],[2,2]以及剩下的 10 个方格中的 1 个里，第 3 种情况，陷阱在[1,3],[3,1]以及剩下的 10 个方格中的 1 个里，总共有 1 + 10 + 10 = 21 种情况;

[1,3]无陷阱可分为 2 种情况，第 1 种，2 个陷阱在[2,2],[3,1]]以及剩下的 10 个方格中的 1 个里，第 2 种情况，陷阱在[2,2]以及剩下的 10 个方格中的 2 个里，总共有 10 + 45 = 55 种情况;

所以，$P(P_{1,3} = True) = 21/(21+55) = 0.276$;

同理，可起得到，当[2,2]中有陷阱的时候，可分为 4 种情况，即[1,3],[3,1]都有陷阱；[1,3]，[3,1]中只有一个有陷阱，[1,3][3,1]都没有陷阱;

共有 1 + 2 * 10 + 45 = 66 种可能赋值;

当[2,2]中没有陷阱的时候，只有一种可能的情况[1,3],[3,1]中都有陷阱，共 10 种可能;

$P(P_{2,2} = True) = 66/(66+10) = 0.868$;

13.24 解：重新计算每个模型的先验概率 P(Frontier);

对于[1,3]有陷阱的的 3 种模型，概率分别为 0.0001,0.0099,0.0099；对于[1,3]无陷阱的 2 种模型，概率分别为 0.0001,0.0099;

$P(P_{1,3} \mid Known,b) =$ 归一化$\langle 0.01 * (0.0001 + 0.0099 + 0.0099),0.99 * (0.0001 + 0.0099)\rangle = \langle 0.1674,0.8326\rangle$;

同理，可得

$P(P_{2,2} \mid Known,b) =$ 归一化$\langle 0.001 * (0.0001 + 0.0099 + 0.0099 + 0.9801), 0.99 * (0.0001)\rangle = \langle 0.9902,0.0098\rangle$;

[3,1]的情况与[1,3]相同，同理可得;

计算结果表明，[2,2]含有陷阱的概率很高，因此概率 Agent 会选择[1,3]或者[3,1]两个方格进入，掉入陷阱的概率为 0.1674;

而逻辑 Agent 会随机选择一个方格进入，掉入陷阱的概率为(0.1674 + 0.1674 + 0.9902) / 3 = 0.4416;

所以，概率 Agent 的性能更好;