

U201714670 范唯 CS1703

朴素贝叶斯作业

请用表 2-1 的训练数据学习一个朴素贝叶斯分类器，表中 $X^{(1)}, X^{(2)}$ 为特征，取值的集合分别为 $A_1 = \{1, 2, 3\}$ ， $A_2 = \{S, M, L\}$ ， Y 为类标记， $Y \in C = \{1, -1\}$ 。并确定 $x = (2, S)^T$ 的类标签 y ，请给出推导具体过程。

表 2-1.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$X^{(1)}$	1	1	1	1	1	2	2	2	2	2	3	3	3	3	3
$X^{(2)}$	S	M	M	S	S	S	M	M	L	L	L	M	M	L	L
Y	-1	-1	1	1	-1	-1	-1	1	1	1	1	1	1	1	-1

解题思路

$$P(Y = 1) = 9/15$$

$$P(Y = -1) = 6/15$$

$$P(X^{(1)} = 1|Y = 1) = 2/9, P(X^{(1)} = 2|Y = 1) = 3/9, P(X^{(1)} = 3|Y = 1) = 4/9$$

$$P(X^{(2)} = S|Y = 1) = 1/9, P(X^{(2)} = M|Y = 1) = 4/9, P(X^{(2)} = L|Y = 1) = 4/9$$

$$P(X^{(1)} = 1|Y = -1) = 3/6, P(X^{(1)} = 2|Y = -1) = 2/6, P(X^{(1)} = 3|Y = -1) = 1/6$$

$$P(X^{(2)} = S|Y = -1) = 3/6, P(X^{(2)} = M|Y = -1) = 2/6, P(X^{(2)} = L|Y = -1) = 1/6$$

对于 $x = (2, S)^T$ 计算：

$$\therefore P(Y = 1|(X^{(1)} = 2 \wedge X^{(2)} = S))$$

$$= P((X^{(1)} = 2 \wedge X^{(2)} = S)|Y = 1)P(Y = 1)/P((X^{(1)} = 2 \wedge X^{(2)} = S))$$

而 $P((X^{(1)} = 2 \wedge X^{(2)} = S))$ 为确定数值，所以可以把这个分母抛弃掉

$$\therefore \text{上式} = P(X^{(1)} = 2|Y = 1) * P(X^{(2)} = S|Y = 1)P(Y = 1)$$

$$\therefore P(Y = 1)P(X^{(1)} = 2|Y = 1)P(X^{(2)} = S|Y = 1) = 1/45$$

$$P(Y = -1)P(X^{(1)} = 2|Y = -1)P(X^{(2)} = S|Y = -1) = 1/15$$

从上述推导过程可以得出 $x = (2, S)^T$ 的类标签 y 应该为 -1.

拉普拉斯平滑(Laplace smoothing): 由于概率计算中间可能会出现所要估计概率值为0的情况，进而使得后验概率的计算结果受到影响，使分类结果产生偏差，所以会采用Laplace校准

$$P_{\lambda}(Y = c_k) = \frac{\sum_{i=1}^N I(y_i=c_k) + \lambda}{N + K\lambda}$$

但是如果需要考虑到拉普拉斯平滑的话。则需要作出以下修正

$$P(Y = 1) = 10/17, P(Y = -1) = 7/17$$

$$P(X^{(1)} = 1|Y = 1) = 3/12, P(X^{(1)} = 2|Y = 1) = 4/12, P(X^{(1)} = 3|Y = 1) = 5/12$$

$$P(X^{(2)} = S|Y = 1) = 2/12, P(X^{(2)} = M|Y = 1) = 5/12, P(X^{(2)} = L|Y = 1) = 5/12$$

$$P(X^{(1)} = 1|Y = -1) = 4/9, P(X^{(1)} = 2|Y = -1) = 3/9, P(X^{(1)} = 3|Y = -1) = 2/9$$

$$P(X^{(2)} = S|Y = -1) = 4/9, P(X^{(2)} = M|Y = -1) = 3/9, P(X^{(2)} = L|Y = -1) = 2/9$$

所以对于给定的 $x = (2, S)T$ 计算：

$$P(Y = 1)P(X^{(1)} = 2|Y = 1)P(X^{(2)} = S|Y = 1) = (10/17)(4/12)(2/12) = 0.0327$$

$$P(Y = -1)P(X^{(1)} = 2|Y = -1)P(X^{(2)} = S|Y = -1) = (7/17)(3/9)(4/9) = 0.0610$$

所以分类结果为 $y = -1$

相关知识

贝叶斯定理： $P(B|A) = P(A|B)P(B)/P(A)$

朴素贝叶斯分类是一种十分简单的分类算法，叫它朴素贝叶斯分类是因为这种方法的思想真的很朴素，朴素贝叶斯的思想基础是这样的：对于给出的待分类项，求解在此项出现的条件下各个类别出现的概率，哪个最大，则认为此待分类项属于哪个类别。通俗来说，就好比这么个道理，你在街上看到一个黑人，我问你你猜这哥们哪里来的，你十有八九猜非洲。为什么呢？因为黑人中非洲人的比率最高，当然人家也可能是美洲人或亚洲人，但在没有其它可用信息下，我们会选择条件概率最大的类别，这就是朴素贝叶斯的思想基础。

$$X = x_1, x_2, \dots, x_m$$

$$Y = y_1, y_2, \dots, y_n$$

$$\text{计算 } P(y_1|X), P(y_2|X), \dots, P(y_n|X)$$

这里找出最大的 $P(y_k|X) X \in y_k$

(我们可以把 X 等想像成迎面走来的穿着 *Air Jordan* 黑人，而 y_1, y_2 可以想象成欧洲人、非洲人等等)

(而 x_1 则是特征黑人， x_2 则是特征穿着 *Air Jordan*)

1. 统计各个类别下各个特征属性的条件概率估计：

$$P(x_1|y_1), P(x_1|y_2), \dots, P(x_m|y_n)$$

2. 并且 $P(y_i|X) = P(X|y_i)P(y_i)/P(X)$

3. 因为各个特征是独立的，所以 $P(X|y_i)P(y_i)$ 可以由下列算式计算

$$P(x_1|y_i)P(x_2|y_i) \dots P(x_m|y_i)P(y_i) = P(y_i) \prod_{j=1}^m P(x_j|y_i)$$

4. 那么可以求得所有的 $P(y_i|X)$ ，得出最大的 $P(y_k|X)$