理论作业-2：朴素贝叶斯法

10225501443 刘蔚璁

Q1: 针对下表的数据，采用拉普拉斯平滑建立贝叶斯分类器，并求点 $x=(3,S)^{\mathrm{T}}$ 的类标记。

	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	S	M	M	S	S	S	M	M	L	L	L	M	M	L
Y	-1	-1	1	1	-1	-1	-1	1	1	1	1	1	1	1	-1

朴素贝叶斯算法

输入：训练数据 $T=\{(x_1,y_1),(x_2,y_2),\ldots,(x_N,y_N)\}$，其中 $x_i=(x_i^{(1)},x_i^{(2)},\ldots,x_i^{(n)})^T$，$x_i^{(j)}$ 是第 i 个样本的第 j 个特征，$x_i^{(j)}\in\{a_{j1},a_{j2},\ldots,a_{jS_j}\}$，$a_{jl}$ 是第 j 个特征可能取的第 l 个值，$j=1,2,\ldots,n,l=1,2,\ldots,S_j,y_i\in\{c_1,c_2,\ldots,c_K\}$；实例 x；

输出：实例 x 的分类。

(1) 计算先验概率及条件概率：

$$P(Y=c_k)=\frac{\sum_{i=1}^N I(y_i=c_k)}{N},\quad k=1,2,\ldots,K$$

$$P(X^{(j)}=a_{jl}\mid Y=c_k)=\frac{\sum_{i=1}^N I(x_i^{(j)}=a_{jl},y_i=c_k)}{\sum_{i=1}^N I(y_i=c_k)}$$

$$j=1,2,\ldots,n;\quad l=1,2,\ldots,S_j;\quad k=1,2,\ldots,K$$

(2) 对于给定的实例 $x=(x^{(1)},x^{(2)},\ldots,x^{(n)})^T$，计算：

$$P(Y=c_k)\prod_{j=1}^n P(X^{(j)}=x^{(j)}\mid Y=c_k),\quad k=1,2,\ldots,K$$

(3) 确定实例 x 的类：

$$y=\arg\max_{c_k} P(Y=c_k)\prod_{j=1}^n P(X^{(j)}=x^{(j)}\mid Y=c_k)$$

拉普拉斯平滑

先验概率

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

条件概率

$$P_\lambda(X^{(j)}=a_{jl}\mid Y=c_k)=\frac{\sum_{i=1}^N I(x_i^{(j)}=a_{jl},y_i=c_k)+\lambda}{\sum_{i=1}^N I(y_i=c_k)+S_j\lambda}$$

其中 K 为类别数，S_j 为某特征的可能取值数。

解

按照拉普拉斯平滑估计概率，取 $\lambda=1$：

$A_1 = \{1, 2, 3\}$, $A_2 = \{S, M, L\}$, $C = \{1, -1\}$。

$P(Y=1) = \frac{10}{17}$, $P(Y=-1) = \frac{7}{17}$

$P(X^{(1)}=1 \mid Y=1) = \frac{3}{12}$, $P(X^{(1)}=2 \mid Y=1) = \frac{4}{12}$, $P(X^{(1)}=3 \mid Y=1) = \frac{5}{12}$

$P(X^{(2)}=S \mid Y=1) = \frac{2}{12}$, $P(X^{(2)}=M \mid Y=1) = \frac{5}{12}$, $P(X^{(2)}=L \mid Y=1) = \frac{5}{12}$

$P(X^{(1)}=1 \mid Y=-1) = \frac{4}{9}$, $P(X^{(1)}=2 \mid Y=-1) = \frac{3}{9}$, $P(X^{(1)}=3 \mid Y=-1) = \frac{2}{9}$

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

对于给定的 $x = (3, S)^T$，计算：

$P(Y=1)P(X^{(1)}=3 \mid Y=1)P(X^{(2)}=S \mid Y=1) = \frac{10}{17} \cdot \frac{5}{12} \cdot \frac{2}{12} = \frac{25}{612} \approx 0.0408$

$P(Y=-1)P(X^{(1)}=3 \mid Y=-1)P(X^{(2)}=S \mid Y=-1) = \frac{7}{17} \cdot \frac{2}{9} \cdot \frac{4}{9} = \frac{56}{1377} \approx 0.0407$

由于 $P(Y=1)P(X^{(1)}=3 \mid Y=1)P(X^{(2)}=S \mid Y=1)$ 更大，所以 $y = 1$。