

多标签分类评价指标介绍（一）

<https://github.com/sangjianshun/Master-School/blob/master/EvaluationIndex1.py>

业务背景

一篇文本

模型一

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.5	0.1	0.15

消极	积极
1	0
0.6	0.4

注意：这个概率和单标签分类之间的区别

一篇文本

模型二

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.6	0.1	0.35

多标签分类评价指标介绍（一）

● 汉明损失 (Hamming Loss)

$$\text{HammingLoss} = \frac{\sum_{i=1}^N \sum_{j=1}^{l_n} Y_{ij} \text{xor} P_{ij}}{N \times l_n}$$

N: 样本数, 这里为1

l_n : 标签数, 这里为5

Y_{ij} : 第*i*条样本是否包含第*j*个标签

P_{ij} : 第*i*个样本的第*j*个标签的预测情况

xor: 异或操作.即不同为1, 相同为0

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.5	0.1	0.15

体育	地理	娱乐	军事	科技
1	0	1	0	0
0	1	1	0	0

$$\text{HammingLoss} = \frac{2}{1 \times 5} = 0.4$$

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.6	0.1	0.35

体育	地理	娱乐	军事	科技
1	0	1	0	0
0	1	1	0	0

$$\text{HammingLoss} = \frac{2}{1 \times 5} = 0.4$$

多标签分类评价指标介绍（一）



覆盖率 (Coverage)

$$Coverage = \frac{\sum_{i=1}^N \max(M_{l_1}, M_{l_2}, \dots, M_{l_n})}{N}$$

N: 样本数, 这里为1

M_{l_n} : 标签 l_n 的预测概率按照所有预测概率从大到小排序的位置, 如果样本不包含标签 l_n 那么 M_{l_n} 会被更新为0

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.5	0.1	0.15

体育	地理	娱乐	军事	科技
1	0	1	0	0
3	2	1	4	5

$$Coverage = \frac{3}{1} = 3$$

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.6	0.1	0.35

体育	地理	娱乐	军事	科技
1	0	1	0	0
4	2	1	5	3

$$Coverage = \frac{4}{1} = 4$$

多标签分类评价指标介绍（一）

1-错误率 (One Error)

$$OneError = \frac{\sum_{i=1}^N I(\arg\max_{l_j} P_{ij} \notin \mathbb{L}_i)}{N}$$

N: 样本数, 这里为1
 $I(\cdot)$: index function, 即统计满足括号里的条件的次数

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.5	0.1	0.15

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.5	0.1	0.15

$$OneError = \frac{0}{1} = 0$$

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.6	0.1	0.35

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.6	0.1	0.35

$$OneError = \frac{0}{1} = 0$$

多标签分类评价指标介绍（一）

排序损失 (Ranking Loss)

$\bar{\mathbb{L}}_i$: 样本不包含的标签集
 \mathbb{L}_i : 样本包含的标签集

$$RankingLoss = \frac{\sum_{j \in \mathbb{L}_i} \sum_{k \in \bar{\mathbb{L}}_i} I(P_{ij} \leq P_{ik})}{|\mathbb{L}_i| * |\bar{\mathbb{L}}_i|}$$

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.5	0.1	0.15

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.5	0.1	0.15

$$RankingLoss = \frac{1}{2*3} = 0.1667$$

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.6	0.1	0.35

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.6	0.1	0.35

$$RankingLoss = \frac{2}{2*3} = 0.3333$$

多标签分类评价指标介绍（一）

● 平均精确率 (Average Precision)

$$Precision = \frac{1}{|\mathbb{L}_i|} \sum_{l_j \in \mathbb{L}_i} \frac{l_j \text{ 在 } P_i \text{ 中的 } \mathbb{L}_i \text{ 子集排名}}{l_j \text{ 在 } P_i \text{ 中的排名}}$$

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.5	0.1	0.15

体育	地理	娱乐	军事	科技
1	0	1	0	0
3	2	1	4	5

$$Precision = \frac{1}{2} \left(\frac{2}{3} + \frac{1}{1} \right) = 0.8333$$

体育	地理	娱乐	军事	科技
1	0	1	0	0
0.3	0.4	0.6	0.1	0.35

体育	地理	娱乐	军事	科技
1	0	1	0	0
4	2	1	5	3

$$Precision = \frac{1}{2} \left(\frac{2}{4} + \frac{1}{1} \right) = 0.75$$