

多标签分类评价指标介绍（二）

<https://github.com/sangjianshun/Master-School/blob/master/algorithm/code/SingleClassEvaluationIndex.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

<https://github.com/sangjianshun/Master-School/tree/master/algorithm/ppt>

多标签分类评价指标介绍（二）



混淆矩阵 (Confusion Matrix)

		标签类别		
预测类别		Positive	Negative	Total
	Positive	TP	FP	预测为正
	Negative	FN	TN	预测为负
	Total	所有正例	所有负例	样本数

1	1	0	0
1	1	0	0
1	0	0	0
1	1	0	1
		0	1

→

TP=3	FP=2
FN=1	TN=4

$$Precision = \frac{TP}{TP + FP} = \frac{3}{5} = 0.6$$

$$Recall = \frac{TP}{TP + FN} = \frac{3}{4} = 0.75$$

$$F_{\alpha} = \frac{(\alpha^2 + 1)P * R}{\alpha^2 * P + R}$$

$$F1\ Score = \frac{2 * P * R}{P + R} = 0.67$$

多标签分类评价指标介绍（二）



多标签场景下

体育

		0	0
1	1	0	0
1	1	0	0
1	0	0	0
1	1	0	1
		0	1

TP=3	FP=2
FN=1	TN=4

地理

		0	0
1	0	0	1
1	1	0	1
1	0	0	0
1	1	0	1
		0	1

TP=2	FP=4
FN=2	TN=2

娱乐

1	0		
1	1	0	0
1	0	0	1
1	1	0	1
1	1	0	0
1	1		

TP=4	FP=2
FN=2	TN=2

军事

1	0		
1	0	0	0
1	0	0	0
1	1	0	1
1	0	0	1
1	1		

TP=2	FP=2
FN=4	TN=2

多标签分类评价指标介绍（二）



F1 Micro和F1 Macro

TP=3	FP=2
FN=1	TN=4

TP=2	FP=4
FN=2	TN=2

TP=4	FP=2
FN=2	TN=2

TP=2	FP=2
FN=4	TN=2

Precision Macro = $\frac{TP}{TP + FP} = \frac{3}{3 + 2} = 0.6$

$\frac{2}{2 + 4} = 0.33$

$\frac{4}{4 + 2} = 0.67$

$\frac{2}{2 + 2} = 0.5$

F1 Macro: 考虑
标签之间的差异。

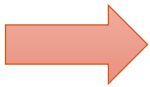
Recall Macro = $\frac{TP}{TP + FN} = \frac{3}{3 + 1} = 0.75$

$\frac{2}{2 + 2} = 0.5$

$\frac{4}{4 + 2} = 0.67$

$\frac{2}{2 + 4} = 0.33$

F1 Macro = $\frac{2 * P * R}{P + R} = 0.67、0.40、0.67、0.40$



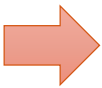
F1 Macro = 0.54

F1 Micro: 不考虑
标签之间的差异。

TP=11	FP=10
FN=9	TN=10

Precision Micro = $\frac{TP}{TP + FP} = \frac{11}{11 + 10} = 0.52$

Recall Micro = $\frac{TP}{TP + FN} = \frac{11}{11 + 9} = 0.55$



F1 Micro = 0.54