

HW5

PB16111485 张劲墩

1.

假设现在有2个检索系统S1和S2，它们在文档集上分别执行查询Q1和Q2，均返回了5个结果，如下表所示：

系统&查询	1	2	3	4	5
S1, Q1	d3	d5	d8	d10	d11
S1, Q2	d1	d2	d7	d11	d13
S2, Q1	d6	d7	d2	d9	d8
S2, Q2	d1	d2	d4	d11	d12

设Q1的相关文档为 {d5, d6, d7, d8}，Q2的相关文档为 {d1, d2, d11, d12, d13} 别计算S1和S2对于查询Q1和Q2的正确率P、召回率R、F1值、P@4和平均正确率AP，并计算S1和S2在所有查询上的MAP值

答：

正确率P: S1, Q1: 2/5; S1, Q2: 4/5; S2, Q1: 3/5; S2, Q2: 4/5;

召回率R: S1, Q1: 1/2; S1, Q2: 4/5; S2, Q1: 3/4; S2, Q2: 4/5;

F1值: S1, Q1: 4/9; S1, Q2: 4/5; S2, Q1: 2/3; S2, Q2: 4/5;

P@4: S1, Q1: 1/2; S1, Q2: 3/4; S2, Q1: 1/2; S2, Q2: 3/4;

AP: S1, Q1: 7/24; S1, Q2: 71/100; S2, Q1: 13/20; S2, Q2: 71/100;

MAP: S1: 0.50083; S2: 1.36;

2.

假设某个检索系统对查询A返回了6个文档，相关度为0-3之间的整数值（3最高，0最低），每个文档与查询的相关度如下所示，请计算查询A的NDCG6值，要求给出计算步骤。

文档	1	2	3	4	5	6
相关度 (0-3)	3	1	3	0	1	2

答：

文档	1	2	3	4	5	6
相关度 (0-3)	3	1	3	0	1	2
DCG	7	7.63093	11.13093	11.13093	11.51778	12.5864
相关度 (0-3)	3	3	2	1	1	0
IDCG	7	11.41651	12.91651	13.34719	13.73404	13.73404
NDCG	1	0.6684	0.86176	0.83395	0.83863	0.916438