

การจับคู่กับ #1

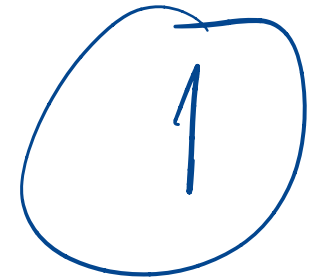
อันดับที่	1	2	3	4	5	6	7	8	9	10
P	1/1	1/2	2/3	2/4	2/5	3/6	3/7	3/8	4/9	5/10
R	1/5	1/5	2/5	2/5	2/5	3/5	3/5	3/5	4/5	5/5

อันดับที่ 10

A										
P	1.0	0.5	0.67	0.5	0.4	0.5	0.43	0.38	0.44	0.5
R	0.2	0.2	0.4	0.4	0.4	0.6	0.6	0.6	0.8	1.0
	<div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> </div>									
	AVGPrec _A = 62.2%									

B										
P	0.0	0.5	0.33	0.25	0.4	0.5	0.57	0.63	0.55	0.5
R	0.0	0.2	0.2	0.2	0.4	0.6	0.6	0.8	1.0	1.0
	<div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> </div>									
	AVGPrec _B = 52%									

AVGPrec_A > AVGPrec_B ∴ A > B



Average Precision

- Often want a single-number effectiveness measure
 - E.g., for a machine-learning algorithm to detect improvement
- Average precision is widely used in IR
- Calculate by averaging precision when recall increases

$$\frac{1.0 + 0.67 + 0.5 + 0.44 + 0.5}{5} =$$

Recall	0.2	0.2	0.4	0.4	0.4	0.6	0.6	0.6	0.8	1.0
Precis.	1.0	0.5	0.67	0.5	0.4	0.5	0.43	0.38	0.44	0.5

พิจารณา Recall ที่เปลี่ยนไป

AvgPrec = 62.2%

Recall	0.0	0.2	0.2	0.2	0.4	0.6	0.8	1.0	1.0	1.0
Precis.	0.0	0.5	0.33	0.25	0.4	0.5	0.57	0.63	0.55	0.5

AvgPrec = 52.0%

$$\frac{0.5 + 0.4 + 0.5 + 0.57 + 0.63}{5}$$

2

R- Precision

① ตรวจประวัติทั้งหมด เทียบ → อธิบายว่านี่
 ② ทดสอบว่าไม่สำคัญ ตรวจประวัติ
 all ที่ผิด

- Precision at the R-th position in the ranking of results for a query that has R relevant documents.

n	doc #	relevant
1	588	x
2	589	x
3	576	
4	590	x
5	986	
6	592	x
7	984	
8	988	
9	578	
10	985	
11	103	
12	591	
13	772	x
14	990	

① ใช้ R = # of relevant docs = 5
 ตรวจประวัติทั้งหมด เทียบ (n)
 คัดที่ n 5 ทน
 ตรวจประวัติ 5

② R-Precision = 3/5 = 0.60

ถ้าตรวจประวัติของข้อมูลทั้งหมด 0/5 = 0
 อธิบายว่านี่

แล้วตรวจว่าตรวจประวัติ
 ที่ผิด

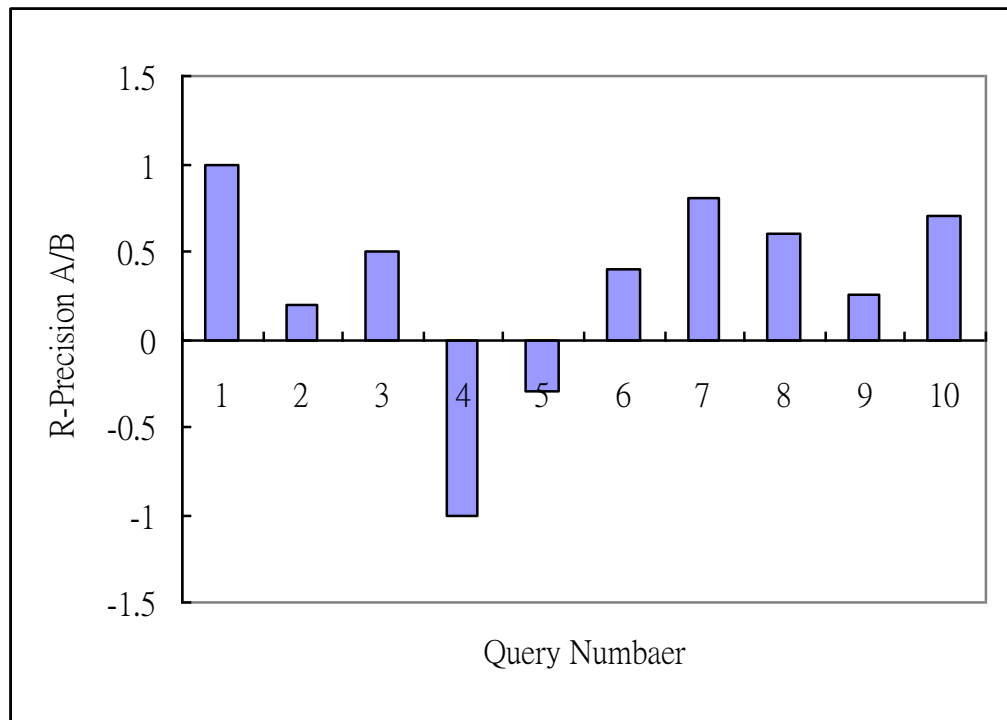
Precision Histograms

ใช้ 2 algo : 10 ตัวอย่าง สังเกตไป แล้ว คำนวณ R-pre เพื่อหาความแตกต่างกับ algo อื่น

- Use R-precision measures to compare the retrieval history of two algorithms through visual inspection

$$RP_{A/B}(i) = RP_A(i) - RP_B(i)$$

เจ็ด + ผลลัพธ์ A
เจ็ด - ผลลัพธ์ B



3

F-Measure

- One measure of performance that takes into account both recall and precision.
- Harmonic mean of recall and precision:

Formula of Harmonic Mean

$$F = \frac{2PR}{P + R} = \frac{2}{\frac{1}{R} + \frac{1}{P}}$$

(Recall) (Precision)

9 ข้อถามสำคัญ:
 ทำไมต้องใช้ค่าเฉลี่ย?
 ทำไมต้องใช้ค่าเฉลี่ย?

- Compared to arithmetic mean, both need to be high for harmonic mean to be high.

4

E Measure (parameterized F Measure)

- A variant of F measure that allows weighting emphasis on precision over recall:

E คือค่าเฉลี่ยถ่วงน้ำหนัก

$$E = \frac{(1 + \beta^2)PR}{\beta^2 P + R} = \frac{(1 + \beta^2)}{\frac{\beta^2}{R} + \frac{1}{P}}$$

ตัวกำหนดถ่วงน้ำหนัก ตัว 100 0-2

เมื่อ $\beta = 1$ คือค่าเฉลี่ย

- Value of β controls trade-off: 9 นี้ 0-1 คือค่าเฉลี่ย
 - $\beta = 1$: Equally weight precision and recall ($E=F$).
 - $\beta > 1$: Weight precision more. มากกว่า
 - $\beta < 1$: Weight recall more. มากกว่า

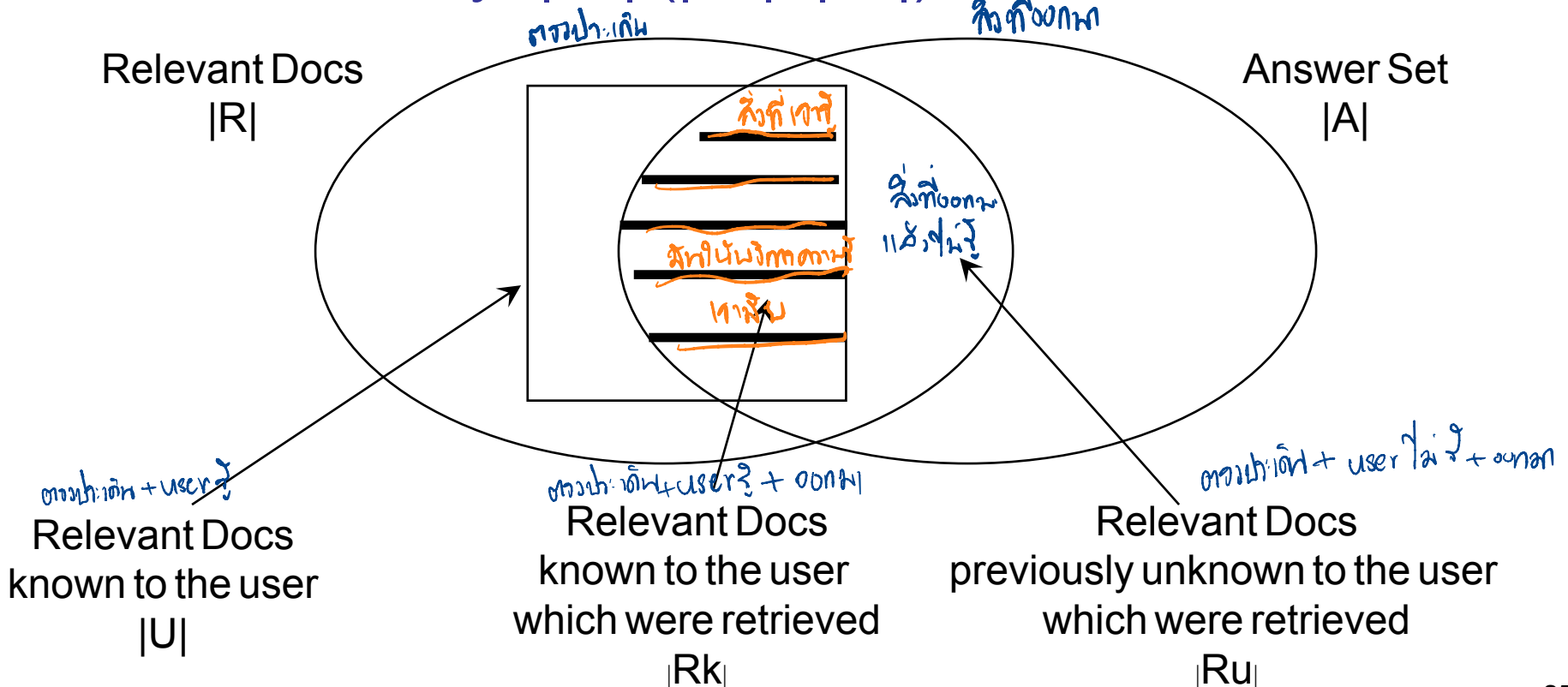
5

User-Oriented Measure

การวัดผล การค้นหา สืบ

– Coverage = $|R_k| / |U|$ ตัวแรก / ตัวทั้งหมด

– Novelty = $|R_u| / (|R_u| + |R_k|)$ ไม่เจอตัวแรก / ตรวจหาสิ่งที่พบที่ออกมา
สิ่งที่ออกมา



Fallout Rate

max
similarity
pre Avg
F
E

- Problems with both precision and recall:
 - Number of irrelevant documents in the collection is not taken into account.
 - Recall is undefined when there is no relevant document in the collection.
 - Precision is undefined when no document is retrieved.

$$\text{Fallout} = \frac{\text{no. of nonrelevant items retrieved}}{\text{total no. of nonrelevant items in the collection}}$$

no. of nonrelevant items retrieved
รวม no. of nonrelevant items in collection ทั้งหมด