

הערכת המודל

Prediction outcome

Actual value

True positive TP	False positive FP
False negative FN	True negative TN

$$\hat{Y} = 0$$

NEGATIVE

$$\hat{Y} = 1$$

POSITIVE

$$Y = 0$$

NOT PREGNANT

TRUE NEGATIVE

You're not pregnant

FALSE POSITIVE

You're pregnant

TYPE 1 ERROR

$$Y = 1$$

PREGNANT

FALSE NEGATIVE

You're not pregnant

TYPE 2 ERROR

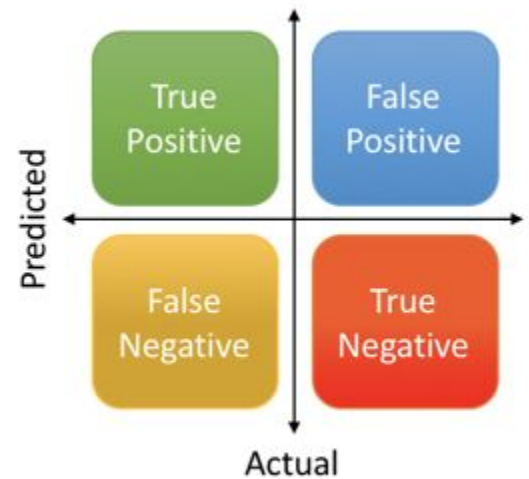
TRUE POSITIVE

You're pregnant

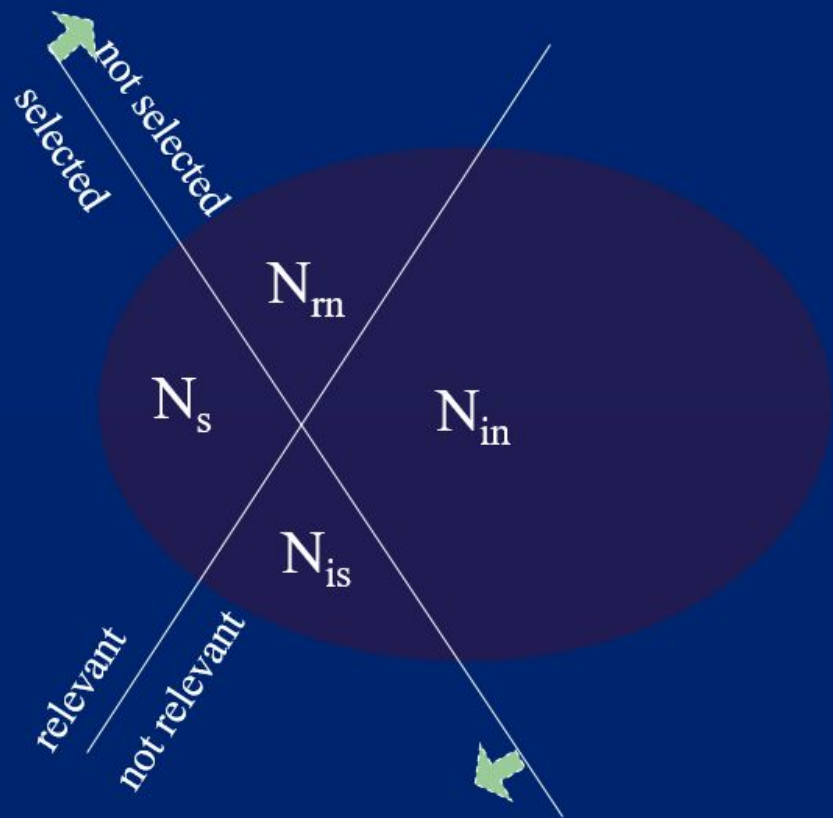
$$\text{Precision} = \frac{\text{True Positive}}{\text{Actual Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{Predicted Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total}}$$



Precision and Recall



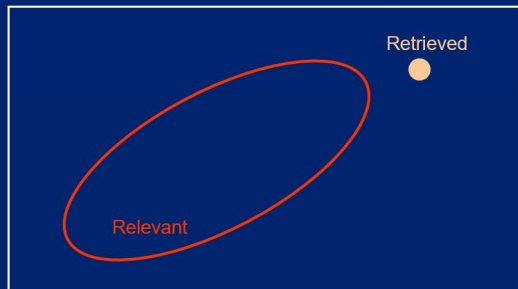
$$\text{Precision} = N_s / (N_s + N_{is})$$

$$\text{Recall} = N_s / (N_s + N_{rn})$$

To improve both P and R
you need to bring the lines
closer together - i.e.
better determination of
relevance.

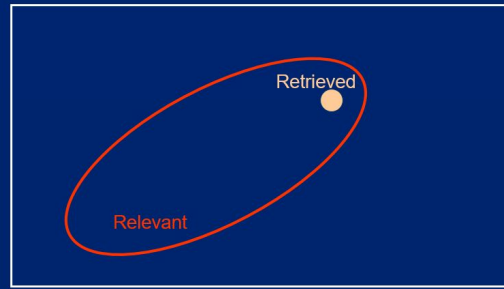
Retrieved vs. Relevant Documents

Very low precision, very low recall (0 in fact)



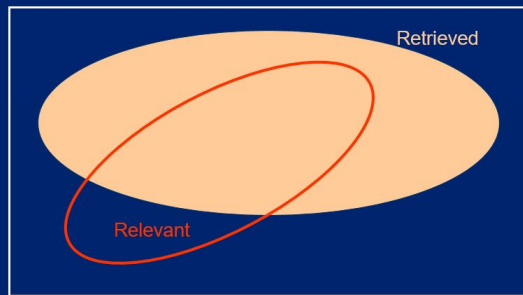
Retrieved vs. Relevant Documents

Very high precision, very low recall



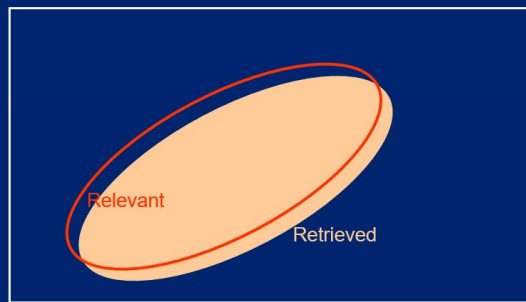
Retrieved vs. Relevant Documents

High recall, low precision



Retrieved vs. Relevant Documents

High precision, high recall (at last!)

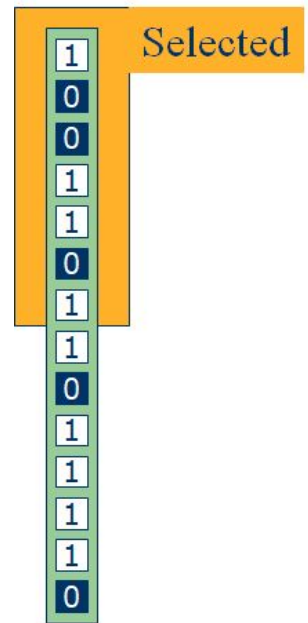


Example – Complete Knowledge

We assume to know the relevance of all the items in the catalogue for a given user

The orange portion is that recommended by the system

Precision= $4/7=0.57$
1's in the orange portion
Recall= $4/9=0.44$
The 1's recommended from the
whole



Example – Incomplete Knowledge

We do not know the relevance of all the items in the catalogue for a given user

The orange portion is that recommended by the system

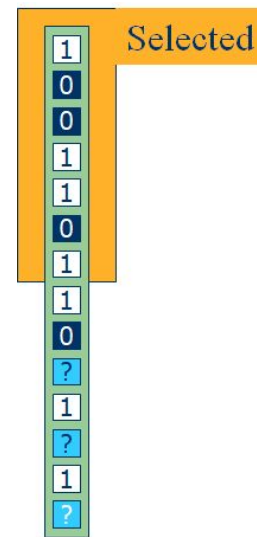
Precision= $4/7=0.57$ – As before

Recall= $4/?$

$4/10 \leq R \leq 4/7$

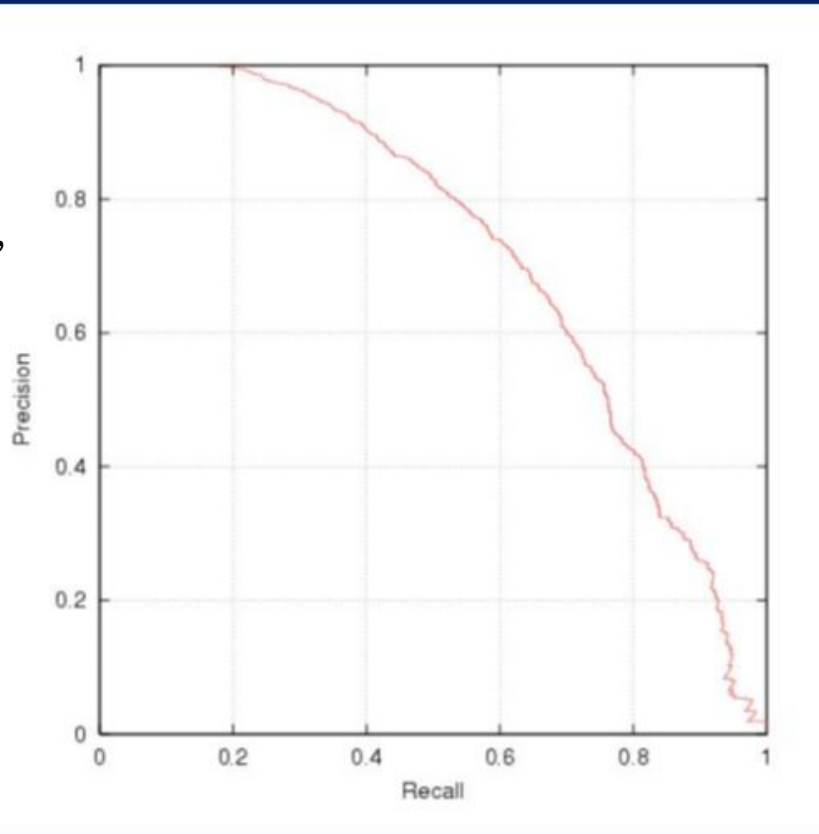
4/10 if all unknown are relevant

4/7 if all unknown are irrelevant



Precision vs. Recall

If you have to recall everything, you will have to keep generating results which are not accurate, hence lowering your precision.

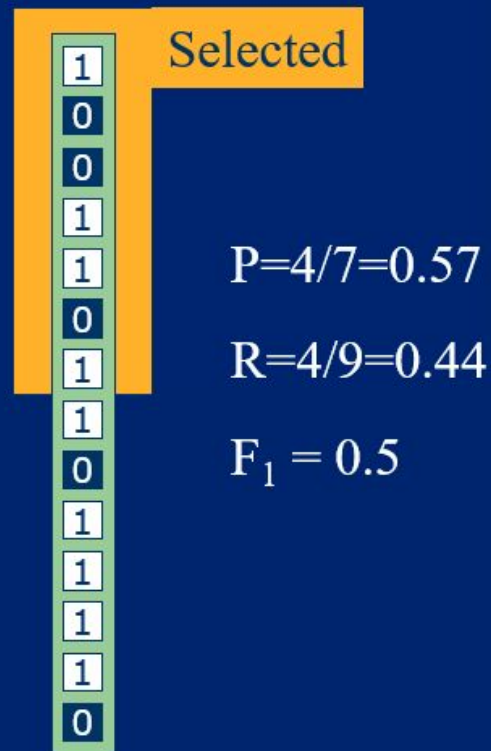


A typical precision and recall curve

F1

- Combinations of Recall and Precision such as F_1
- Typically systems with high recall have low precision and vice versa

$$F_1 = \frac{2PR}{P+R}$$



Consider a documents collection made of 100 documents.

Given a query q , the set of documents relevant to the users is $D^* = \{d_3, d_{12}, d_{34}, d_{56}, d_{98}\}$. An IR system retrieves the following documents $D = \{d_3, d_{12}, d_{35}, d_{56}, d_{66}, d_{88}, d_{95}\}$

- (a) Compute the number of True-Positives, True-Negatives, False-Positives, False-Negatives
- (b) Compute Precision, Recall, Balanced F-measure, Accuracy

Compute the number of True-Positives, True-Negatives, False-Positives, False-Negatives

$$TP = 3$$

$$FP = 4$$

$$FN = 2$$

$$TN = 91$$

Compute Precision, Recall, Balanced F-measure, Accuracy

$$P = \frac{3}{7}$$

$$R = \frac{3}{5}$$

$$F = \frac{1}{2}$$

$$A = \frac{94}{100}$$

An IR system produces the following rankings in answer to queries q_1 and q_2 . The underscored documents are the ones relevant to the user.

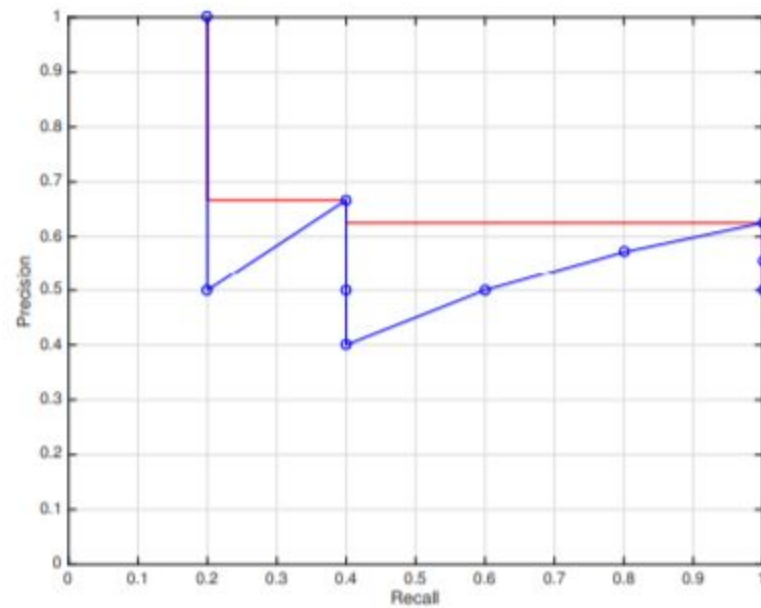
R	q_1	q_2
1	<u>A</u>	<u>F</u>
2	L	<u>G</u>
3	<u>G</u>	D
4	F	<u>E</u>
5	D	L
6	<u>E</u>	I
7	<u>B</u>	H
8	<u>H</u>	C
9	I	<u>B</u>
10	C	A

- (a) Draw the precision-recall curve and the interpolated precision-recall curve
- (b) Compute the Mean Average Precision
- (c) Compute the R-precision
- (d) Draw the Receiver-Operating-Characteristic

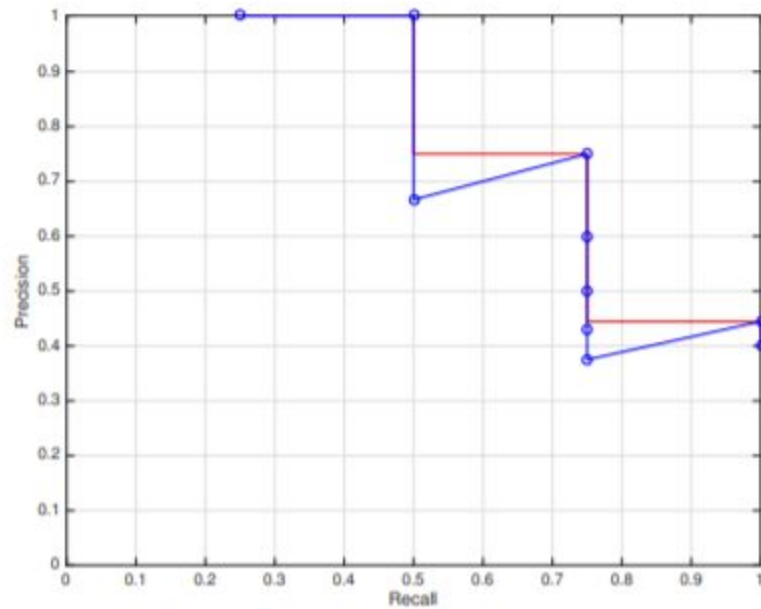
Draw the precision-recall curve and the interpolated precision-recall curve

Retrieved documents	P_{q_1}	R_{q_1}	P_{q_2}	R_{q_2}
1	1/1	1/5	1/1	1/4
2	1/2	1/5	2/2	2/4
3	2/3	2/5	2/3	2/4
4	2/4	2/5	3/4	3/4
5	2/5	2/5	3/5	3/4
6	3/6	3/5	3/6	3/4
7	4/7	4/5	3/7	3/4
8	5/8	5/5	3/8	3/4
9	5/9	5/5	4/9	4/4
10	5/10	5/5	4/10	4/4

Precision-Recall for q_1



Precision-Recall for q_2



Compute the Mean Average Precision

$$AP_{q_1} = \frac{1/1 + 2/3 + 3/6 + 4/7 + 5/8}{5} = 0.67$$

$$AP_{q_2} = \frac{1/1 + 2/2 + 3/4 + 4/9}{4} = 0.80$$

$$MAP = \frac{0.67 + 0.80}{2} = 0.74$$

Compute the R-precision

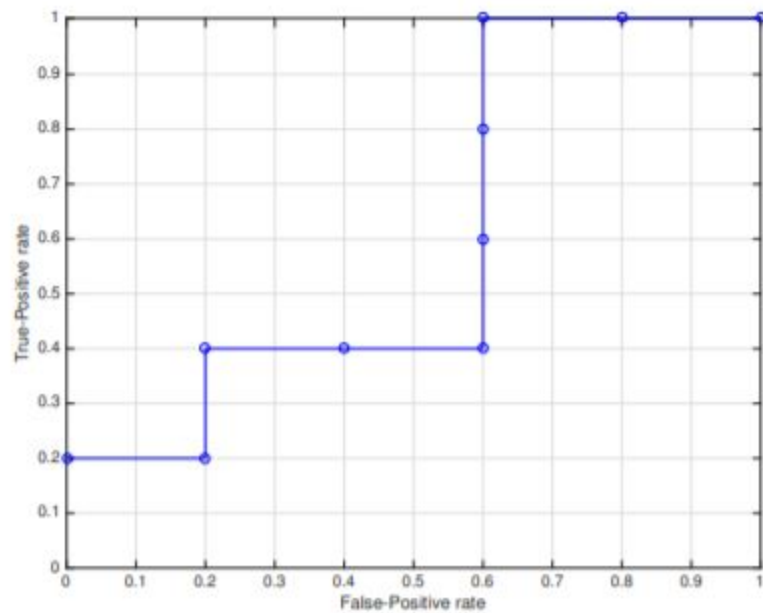
$$Rp_{q_1} = 2/5$$

$$Rp_{q_2} = 3/4$$

Draw the Receiver-Operating-Characteristic

Retrieved documents	FPr_{q_1}	TPr_{q_1}	FPr_{q_2}	TPr_{q_2}
1	0/5	1/5	0/6	1/4
2	1/5	1/5	0/6	2/4
3	1/5	2/5	1/6	2/4
4	2/5	2/5	1/6	3/4
5	3/5	2/5	2/6	3/4
6	3/5	3/5	3/6	3/4
7	3/5	4/5	4/6	3/4
8	3/5	5/5	5/6	3/4
9	4/5	5/5	5/6	4/4
10	5/5	5/5	6/6	4/4

ROC for q_1



ROC for q_2

