

PRÉSENTATION DES MÉTRIQUES ‘MAP’ ET ‘NDCG’

G8 Innovation | 10/01/2020



VUE D'ENSEMBLE

Evaluer un SIR

Mesures d'évaluations

MAP

NDCG

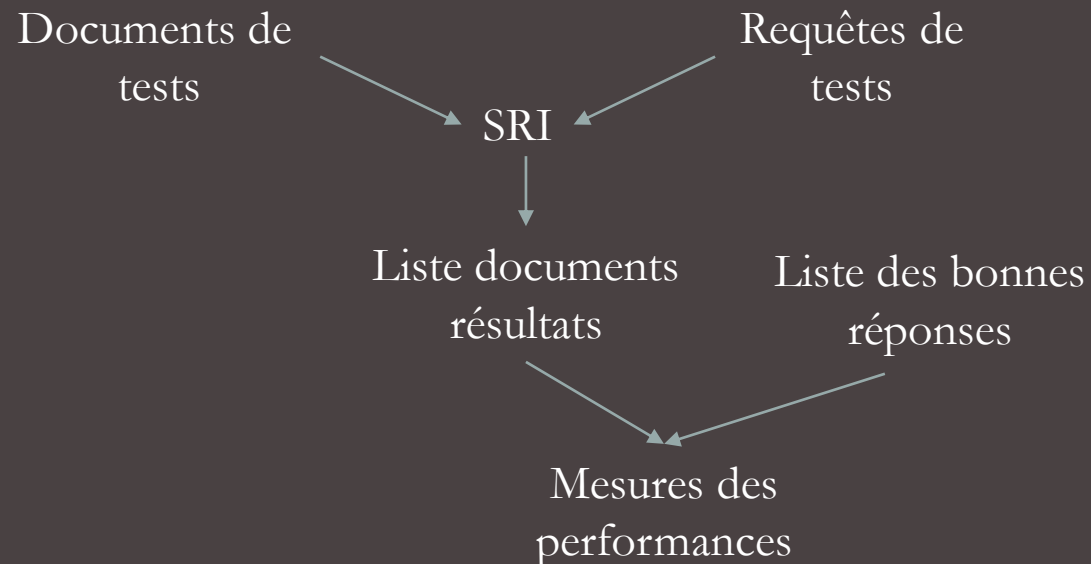


Comment évaluer un système de recherches d'informations ?

* Que faut-il ?

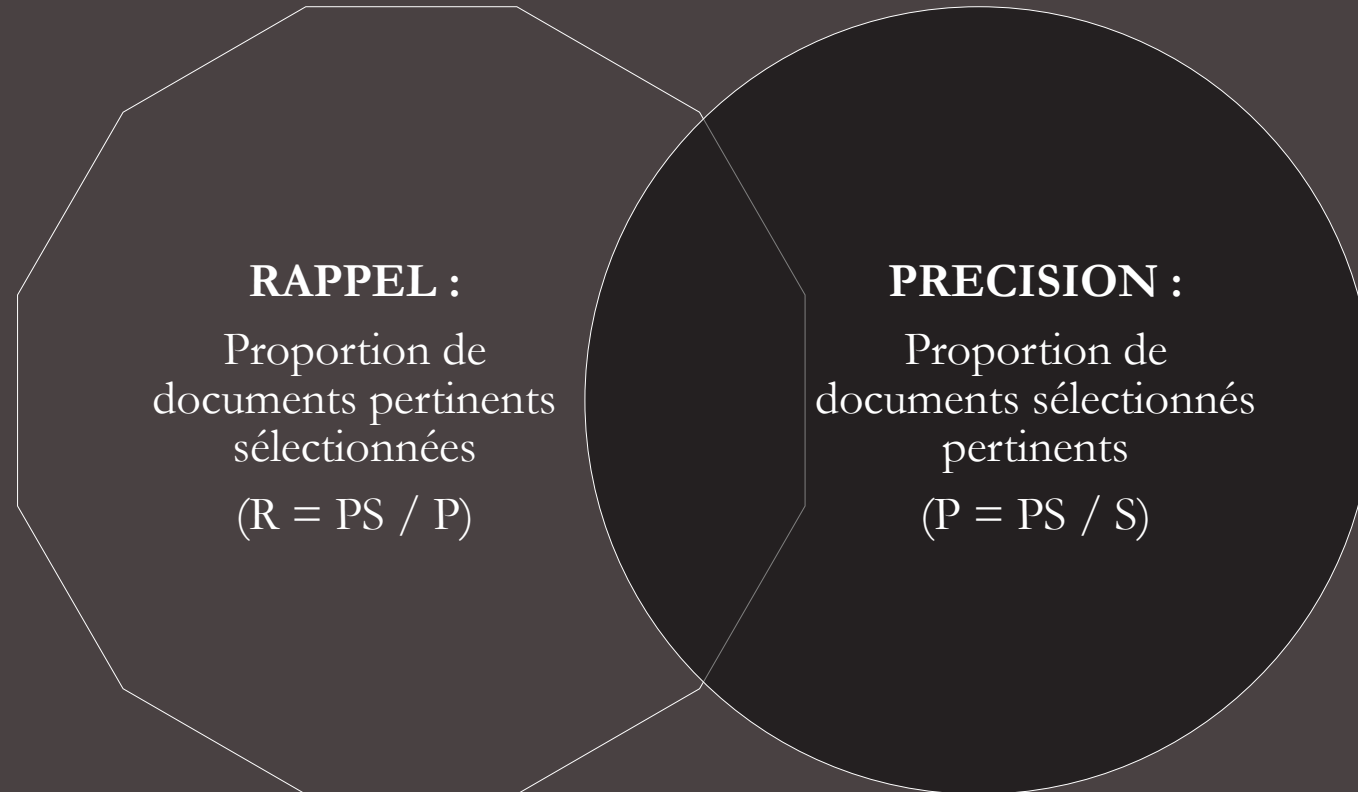
- Un corpus de documents tests à interroger
- Des requêtes de test (minimum 25)
- Jugements de pertinences (correspond aux documents qu'aurait dû retourner le système, pas ceux qu'il a en réalité retournés)

* Démarche d'évaluation :



MESURES D'EVALUATIONS

Deux notions essentielles :



P = Nombre de documents Pertinents

S = Nombre de documents Sélectionnés

PS = Nombre de documents Sélectionnées et qui sont Pertinents

MAP

MESURES AGREGÉES

- Pour une requête : Average Precision (moyenne des précisions calculées pour chaque point de rappel pour une requête).

$$AP(q) = \frac{1}{R_q} \sum_r I(r) * Prec_q(r)$$

- Avec :
- r le rang du document
 - R_q le nombre total de documents pertinents pour la requête q
 - $I(r)$ vaut 1 si le document est pertinent 0 sinon
 - Prec la précision (voir diapo précédente).

- Pour plusieurs requêtes : Mean Average Precision

$$MAP = \frac{1}{|Q|} \sum_{q \in Q} AP(q)$$

- Avec :
- Q l'ensemble des requêtes
 - $AP(q)$ = Average Precision pour une requête

EXEMPLE - MAP

```
[14] # Résultats du système d'informations
res = {
  'q1': {
    'd1': 1.0,
    'd2': -0.1,
    'd3': 1.5,
  },
  'q2': {
    'd1': 1.5,
    'd2': 0.2,
    'd3': 0.5,
  }
}

# Documents "réellement" pertinents
qrels = {
  'q1': {
    'd2': 1
  },
  'q2': {
    'd2': 1,
    'd3': 1,
  },
}

# Evaluation -> métrique 'MAP'
print("Par requête :\n", pt.Utls.evaluate(res, qrels, metrics = ['map'], perquery=True),
      "\nMoyenne :\n", pt.Utls.evaluate(res, qrels, metrics = ['map']))

# Evaluation -> métrique 'NDCG'
print("\nPar requête :\n", pt.Utls.evaluate(res, qrels, metrics = ['ndcg'], perquery=True),
      "\nMoyenne :\n", pt.Utls.evaluate(res, qrels, metrics = ['ndcg']))

Par requête :
{'q1': {'map': 0.3333333333333333}, 'q2': {'map': 0.5833333333333333}}
Moyenne :
{'map': 0.4583333333333333}

Par requête :
{'q1': {'ndcg': 0.5}, 'q2': {'ndcg': 0.6934264036172708}}
Moyenne :
{'ndcg': 0.5967132018086354}
```

- Ensemble de documents : $D = \{D1, D2, D3\}$
- Ensemble des requêtes : $Q = \{Q1, Q2\}$
- Résultats du système de recherche d'information :

Q1	Q2
D3 (1.5)	D1 (1.5)
D1 (1.0)	D3 (0,5)
D2 (-0,1)	D2 (0,2)

- Jugements de pertinences :

Q1	Q2
D1 – NP (0)	D1 – NP (0)
D2 – P (1)	D2 – P (1)
D3 – NP (0)	D3 – P (1)

- Résumé :

RANG DOC	DOC	PERTINENCE	PRESICION (P)	DOC	PERTINEN CE	PRESICION (P)
	Q1			Q2		
1	D3	NP	0/1=0	D1	NP	0/1=0
2	D1	NP	0/2=0	D3	P	1/2=0,5
3	D2	P	1/3=0.33	D2	P	2/3=0.67

- $AP(Q1) = \frac{1}{1} \left(1 * \frac{1}{3} \right) = 0,33$
- $AP(Q2) = \frac{1}{2} \left(1 * \frac{1}{2} + 1 * \frac{2}{3} \right) = 0,5833$

}

$MAP = \frac{1}{2} (0,33 * 0,5833) = 0,4583$

NDCG

MESURES ADAPTEES À UNE MESURE GRADUELLE DE LA PERTINENCE

- Pertinence graduelle : Très Pertinent (2), Pertinent (1), Non Pertinent (0) => Labels

- Discounted Cumulative Gain (DCG) :

$$DCG(Q) = \sum_{r=1}^N P_u(r) * Utilité(r)$$

Avec : - $Utilité(r) = 2^{pertinence} - 1$

- Pertinence : pertinence graduelle

- $P_u(r)$: probabilité que l'utilisateur visite le document à la position $r = \frac{1}{\log_2(r+1)}$

- Normalized Discounted Cumulative Gain (NDCG) :

$$NDCG(Q) = \frac{DCG(Q)}{OptDCG}$$

Avec : - OptDCG : valeur DCG optimale

EXEMPLE

- Résultats :

Q1	Q2
D3 (1.5)	D1 (1.5)
D1 (1.0)	D3 (0,5)
D2 (-0,1)	D2 (0,2)

- Jugements de pertinence :

Q1	Q2
D1 – NP (0)	D1 – NP (0)
D2 – P (1)	D2 – P (1)
D3 – NP (0)	D3 – P (1)

- DCG :

RANG DOC	Q1				Q2			
	Doc	Pertinence	Utilité(r)	Utilité(r)*Pu(r)	Doc	Pertinence	Utilité(r)	Utilité(r)*Pu(r)
1	D3	0 (NP)	0	0/log2(2)=0	D1	0 (NP)	1	0/log2(2)=0
2	D1	0 (NP)	0	0/log2(3)=0	D3	1 (P)	1	1/log2(3)=0.63
3	D2	1 (P)	1	1/log2(4)=0,5	D2	1 (P)	1	1/log2(4)=0.5

$$DCG(Q1) = 1 * \log_2(4) = 0,5$$

$$DCG(Q2) = 1 * \log_2(3) + 1 * \log_2(4) = 1,13$$

Par requête :
 {'q1': {'ndcg': 0.5}, 'q2': {'ndcg': 0.6934264036172708}}
 Moyenne :
 {'ndcg': 0.5967132018086354}

$$NDCG(Q1) = \frac{0,5}{1} = 0,5$$

$$NDCG(Q2) = \frac{1,13}{1,63} = 0,693$$

- OptDCG :

RANG DOC	Q1				Q2			
	Doc	Pertinence	Utilité(r)	Utilité(r)*Pu(r)	Doc	Pertinence	Utilité(r)	Utilité(r)*Pu(r)
1	D2	1 (P)	1	1/log2(2)=1	D2	1 (P)	1	1/log2(2)=1
2	D1	0 (NP)	0	0/log2(3)=0	D3	1 (P)	1	1/log2(3)=0.63
3	D3	0 (NP)	0	0/log2(4)=0	D1	0 (NP)	0	0/log2(4)=0

$$OptDCG(Q1) = 1 * \log_2(2) = 1$$

$$OptDCG(Q2) = 1 * \log_2(2) + 1 * \log_2(3) = 1,63$$

NDCG À VALEUR
SEUIL...