

# MESURES D'ÉVALUATIONS

## Accès et Recherche d'Information

Supposons maintenant que les documents  $d_i \in D$ , soient triés dans l'ordre décroissant de leur mesure de pertinence pour la requête  $q$ . En découpant l'ensemble de ces documents triés en  $k$  sous ensembles  $D_k(q) = \{d_{j(1)}, \dots, d_{j(k)}\}$  et en calculant les mesures de précision/rappel sur chacun de ces  $k$  sous-ensembles on obtient *le graphe de précision et de rappel* qui indique comment les documents pertinents et non-pertinents sont répartis parmi les documents triés.

$$R@k = \frac{|D_k^{per}(q)|}{|D^{per}(q)|}$$
$$P@k = \frac{|D_k^{per}(q)|}{|D_k(q)|}$$

Où  $D^{per}(q)$  est l'ensemble de tous les documents pertinents pour la requête  $q$  et  $D_r^{per}(q) = D^{per}(q) \cap D_k(q)$  est l'ensemble des documents pertinents retournés par le système parmi les  $q$  premiers documents les plus similaires.

Le rappel mesure quelle est la proportion de documents pertinents retournés par le système et la précision le pourcentage de documents pertinents parmi ceux retournés.

1 COURBE DE PRÉCISION RAPPEL
------------------------------

Pour une collection constituée de 8 documents, on suppose qu'il y a 4 documents pertinents pour la requête  $q$  notés par un + et 4 autres non-pertinents notés par un -. Le tableau ci-dessous donne pour chaque document  $d_i$ ,  $i \in \{1, \dots, 8\}$  sa pertinence pour la requête  $q$  ainsi que la réponse du système. Dans ce tableau, la colonne 2 indique les jugements de pertinence par rapport à  $q$  et la colonne 3 la réponse du système triée dans l'ordre décroissante.

$d_i$	pertinent pour la requête $q$	Réponse du système
1	+	0.95
2	-	0.83
3	-	0.81
4	+	0.76
5	-	0.56
6	-	0.45
7	+	0.32
8	+	0.21

1.1 Compléter le tableau de précision et de rappel ci-dessous.

$d_i$	pertinent pour la requête $q$	$R@k$	$P@k$
1	+		
2	-		
3	-		
4	+		
5	-		
6	-		
7	+		
8	+		

1.2 Quelle est la mesure AvP de ce système?

1.3 Construire la courbe de précision en fonction du rappel.

Pour lisser la courbe obtenue, il existe une technique basée sur la moyenne statistique. Cette technique consiste à estimer une fonction qui assigne à chaque valeur du rappel  $\rho \in [0, 1]$  une valeur  $\Pi(\rho)$  de précision telle que :

$$\Pi(\rho) = \max \{P@k \mid R@k \geq \rho\}$$

La fonction  $\Pi_k(\rho)$  peut être interprétée comme le pourcentage de documents pertinents qu'un utilisateur observera s'il veut atteindre un taux de rappel d'au moins  $\rho$ . En pratique, les courbes de précision/rappel sont obtenues en calculant  $\Pi(\rho)$  pour les 11 valeurs de  $\rho \in \{0, 0.1, \dots, 1.0\}$

1.4 Lisser la courbe obtenue suivant la technique de la moyenne statistique.

2 COMPARAISON ENTRE SYSTÈMES
------------------------------

Soit  $\mathcal{S}_1$  et  $\mathcal{S}_2$  deux systèmes de recherche qui pour une requête  $q$  donnée renvoient les listes de documents triées suivantes :

$k$	$\mathcal{S}_1$	$\mathcal{S}_2$
1	-	+
2	+	+
3	+	-
4	-	+
5	+	-
6	-	-
7	+	-
8	-	+
9	-	-
10	-	-

- 2.1 Quelles sont les  $P@k$  et  $R@k$  de ces deux systèmes?
- 2.2 En déduire les AvP de ces deux systèmes ainsi que leurs courbes de précision/rappel
- 2.3 Au vue de ces résultats, quel est le système le plus performant et pourquoi?