

Λύσεις 1^{ης} Σειράς Ασκήσεων
 (Αξιολόγηση της Αποτελεσματικότητας της Ανάκτησης)

Άσκηση 1

Θεωρείστε μια συλλογή αξιολόγησης που αποτελείται από 20 έγγραφα $\{d_1, \dots, d_{20}\}$. Η συλλογή αξιολόγησης περιλαμβάνει μια επερώτηση q ; για την οποία γνωρίζουμε ότι τα έγγραφα της συλλογής που είναι συναφή με αυτήν είναι 5, συγκεκριμένα τα $\{d_1, d_{11}, d_{18}, d_{19}, d_{20}\}$. Θέλουμε να αξιολογήσουμε την αποτελεσματικότητα τριών συστημάτων $S1, S2$ και $S3$. Για το λόγο αυτό υποβάλλουμε σε κάθε σύστημα την επερώτηση q και λαμβάνουμε τις εξής απαντήσεις:

$$\begin{aligned} \text{Ans}(S1, q) &= \langle d_{11}, d_4, d_{18}, d_2, d_9, d_7, d_8, d_6, d_1, d_{20} \rangle \\ \text{Ans}(S2, q) &= \langle d_9, d_7, d_5, d_6, d_{11}, d_4, d_8, d_2, d_1, d_{18}, d_{19}, d_{20} \rangle \\ \text{Ans}(S3, q) &= \langle d_{18}, d_{11}, d_1, d_5, d_2 \rangle \end{aligned}$$

Το αριστερότερο στοιχείο της κάθε απάντησης παριστάνει το υψηλότερα διαβαθμισμένο έγγραφο, αυτό που το σύστημα υπολόγισε ως το πιο συναφές με την επερώτηση q . Συγκρίνετε τα τρία αυτά συστήματα ως προς τα εξής μέτρα: (α) Ακρίβεια (Precision). (β) Ανάκληση (Recall), (γ) F-Measure, (δ) R-Ακρίβεια (R-Precision) και (ε) Fallout. Δώστε τα συγκεντρωτικά στοιχεία σε ένα πίνακα της μορφής:

System	Precision	Recall	F-Measure	R-Precision	Fallout
S1					
S2					
S3					

Ranking of Systems					
--------------------	--	--	--	--	--

Σχολιάστε το αποτέλεσμα της σύγκρισης.

Λύση

(α) Precision

- Το $S1$ επιστρέφει συνολικά 10 έγγραφα εκ των οποίων τα 4 είναι συναφή επομένως

$$\text{Precision}(S1) = \frac{4}{10} = 0.4$$
- Το $S2$ επιστρέφει συνολικά 12 έγγραφα εκ των οποίων τα 5 είναι συναφή επομένως

$$\text{Precision}(S2) = \frac{5}{12} = 0.417$$
- Το $S3$ επιστρέφει συνολικά 5 έγγραφα εκ των οποίων τα 3 είναι συναφή επομένως

$$\text{Precision}(S3) = \frac{3}{5} = 0.6$$

(β) Recall

Το πλήθος των συνολικών συναφών εγγράφων για την επερώτηση q είναι 5.

- Το $S1$ επιστρέφει 4 συναφή, επομένως $Recall(S1) = \frac{4}{5} = 0.8$
- Το $S2$ επιστρέφει 5 συναφή, επομένως $Recall(S2) = \frac{5}{5} = 1$
- Το $S3$ επιστρέφει 3 συναφή, επομένως $Recall(S3) = \frac{3}{5} = 0.6$

(γ) F-Measure

Ως F-Measure ορίζεται το αρμονικό μέσο της ανάκλησης και της ακρίβειας. Συγκεκριμένα :

$$F - Measure = \frac{2 \cdot P \cdot R}{P + R} = \frac{2}{\frac{1}{P} + \frac{1}{R}}$$

- $F - Measure(S1) = \frac{2 \cdot P(S1) \cdot R(S1)}{P(S1) + R(S1)} = \frac{2 \cdot 0.4 \cdot 0.8}{0.4 + 0.8} = \frac{0.64}{1.2} = 0.533$
- $F - Measure(S2) = \frac{2 \cdot P(S2) \cdot R(S2)}{P(S2) + R(S2)} = \frac{2 \cdot 0.417 \cdot 1}{0.417 + 1} = \frac{0.834}{1.417} = 0.589$
- $F - Measure(S3) = \frac{2 \cdot P(S3) \cdot R(S3)}{P(S3) + R(S3)} = \frac{2 \cdot 0.6 \cdot 0.6}{0.6 + 0.6} = \frac{0.72}{1.2} = 0.6$

(δ) R-Precision

R-Precision είναι η ακρίβεια ενός συστήματος στην R-θέση της διάταξης των αποτελεσμάτων όπου R είναι το πλήθος των συναφών εγγράφων για μια επερώτηση.

Εδώ το συνολικό πλήθος των συναφών εγγράφων για την επερώτηση q είναι 5. ($R=5$)

- $R - Precision(S1) = Precision(S1)_{position=5} = \frac{2}{5} = 0.4$
- $R - Precision(S2) = Precision(S2)_{position=5} = \frac{1}{5} = 0.2$
- $R - Precision(S3) = Precision(S3)_{position=5} = \frac{3}{5} = 0.6$

(ε) Fallout

Το Fallout είναι το πλήθος των μη συναφών εγγράφων που έχει ανακτήσει το σύστημα για μία επερώτηση q προς το συνολικό πλήθος των μη συναφών εγγράφων .

Το συνολικό πλήθος των μη συναφών εγγράφων για την επερώτηση q είναι 15.

- Το $S1$ επιστρέφει 6 μη συναφή έγγραφα, άρα $Fallout(S1) = \frac{6}{15} = 0.4$
- Το $S2$ επιστρέφει 7 μη συναφή έγγραφα, άρα $Fallout(S2) = \frac{7}{15} = 0.467$
- Το $S3$ επιστρέφει 2 μη συναφή έγγραφα, άρα $Fallout(S3) = \frac{2}{15} = 0.133$

Συγκεντρωτικά ο πίνακας έχει ως εξής :

System	Precision	Recall	F-Measure	R-Precision	Fallout
S1	0.4	0.8	0.533	0.4	0.4
S2	0.417	1	0.589	0.2	0.467
S3	0.6	0.6	0.6	0.6	0.133

Ranking of Systems	<S3,S2,S1>	<S2,S1,S3>	<S3,S2,S1>	<S3,S1,S2>	<S3,S1,S2>
--------------------	------------	------------	------------	------------	------------

Precision: Και τα 3 συστήματα έχουν τιμές ακρίβειας πολύ μικρότερες από 1. Αυτό σημαίνει ότι και τα 3 συστήματα επιστρέφουν εκτός από συναφή έγγραφα και αρκετά μη-συναφή. Ωστόσο το τρίτο σύστημα υπερέρχει επειδή το 60 % των εγγράφων που περιέχει είναι συναφή. Από την άλλη τα άλλα 2 επιστρέφουν περισσότερα μη συναφή έγγραφα. Τυπικά το S2 είναι καλύτερο από το S1. Οι τιμές είναι λογικές αν αναλογιστούμε ότι το S3 επιστρέφει μόνο 3 συναφή έγγραφα (σε αντίθεση με 4, 5 για τα S1, S2 αντίστοιχα) όμως επιστρέφει μόνο 2 μη συναφή.

Recall: Όπως φαίνεται και από τις τιμές της ανάκλησης του πίνακα το καλύτερο σύστημα όσον αφορά την ανάκληση είναι το S2 το οποίο επιστρέφει όλα τα συναφή έγγραφα της συλλογής. Ακολουθεί το S1 και τελευταίο είναι το S3.

F-Measure: Χρησιμοποιώντας το *F-Measure* μπορούμε να προσδιορίσουμε ποιο σύστημα υπερέρχει κοιτάζοντας το και τόσο από την σκοπιά της ακρίβειας όσο και από την σκοπιά της ανάκλησης. Αυτό διότι ένα σύστημα έχει υψηλό *F-Measure* όταν έχει υψηλό *Recall* και υψηλό *Precision*. Έτσι το S3 είναι προτιμητέο επειδή συνδυάζει παράλληλα υψηλές τιμές για *Recall* και *Precision*. Ακολουθεί το S2 και τελευταίο είναι το S1.

R-Precision: Το Καλύτερο σύστημα από πλευράς R-ακρίβειας είναι το S3 επειδή επιστρέφει όλα τα συναφή έγγραφα που έχει ανακτήσει στις 5 πρώτες θέσεις. Ακολουθεί το S1 και τέλος το S2. Παρατηρούμε εδώ ότι αν και τα S1 και S2 επιστρέφουν περισσότερα συναφή έγγραφα από το S3, εντούτοις δεν τα επιστρέφουν στην αρχή της διάταξης. Το S3 αντίθετα επιστρέφει πρώτα τα συναφή έγγραφα και γι' αυτό είναι προτιμότερο από πλευράς R-ακρίβειας.

Fallout: Το S3 είναι το σύστημα το οποίο επιστρέφει τα λιγότερα μη-συναφή έγγραφα και άρα είναι προτιμότερο από πλευράς *Fallout*. Ακολουθεί το S1 και μετά το S2.

Άσκηση 2

Σχεδιάστε τις καμπύλες ακρίβειας/ανάκλησης (P/R curves) των συστημάτων της προηγούμενης άσκησης. Για κάθε σύστημα δώστε 2 γραφήματα: ένα που να απεικονίζει τα P/R σημεία όπως προκύπτουν από τις απαντήσεις και ένα χρησιμοποιώντας κανονικοποιημένα επίπεδα ανάκλησης (standard recall levels). Αν βλέπατε μόνο αυτά τα 2 γραφήματα (και όχι τις απαντήσεις) θα μπορούσατε να επιλέξετε το καλύτερο σύστημα;

Ένας εναλλακτικός τρόπος αξιολόγησης της αποτελεσματικότητας ενός συστήματος είναι οι καμπύλες *Recall-Fallout*. Ορίζονται ανάλογα με τις καμπύλες *Precision-Recall*, μόνο που τώρα ο άξονας X έχει τιμές *Fallout*, ενώ ο Y έχει τιμές *Recall*. Σχεδιάστε τις καμπύλες *Recall-Fallout* των συστημάτων της προηγούμενης άσκησης.

Λύση

Precision/Recall Curves

S1

Υπολογίζουμε τα σημεία για κάθε συναφές έγγραφο στα αποτελέσματα του συστήματος. Το S1 ανακτά συνολικά 4 έγγραφα

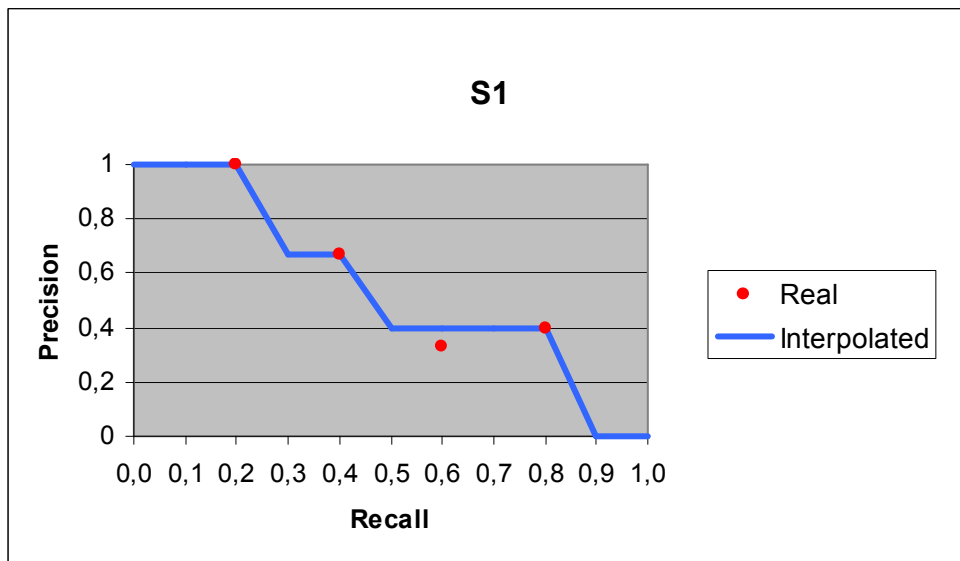
$$\text{1}^\circ \text{ Συναφές: } \text{Recall}(S1) = \frac{1}{5} = 0.2 \quad \text{Precision}(S1) = \frac{1}{1} = 1$$

2° Συναφές: $Recall(S1) = \frac{2}{5} = 0.4$ $Precision(S1) = \frac{2}{3} = 0.667$

3° Συναφές: $Recall(S1) = \frac{3}{5} = 0.6$ $Precision(S1) = \frac{3}{9} = 0.333$

4° Συναφές: $Recall(S1) = \frac{4}{5} = 0.8$ $Precision(S1) = \frac{4}{10} = 0.4$

Επίσης χρησιμοποιούμε τα καθιερωμένα επίπεδα ανάκλησης
 $\{0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$



S2

Το S2 ανακτά συνολικά 5 έγγραφα

1° Συναφές: $Recall(S2) = \frac{1}{5} = 0.2$ $Precision(S2) = \frac{1}{5} = 0.2$

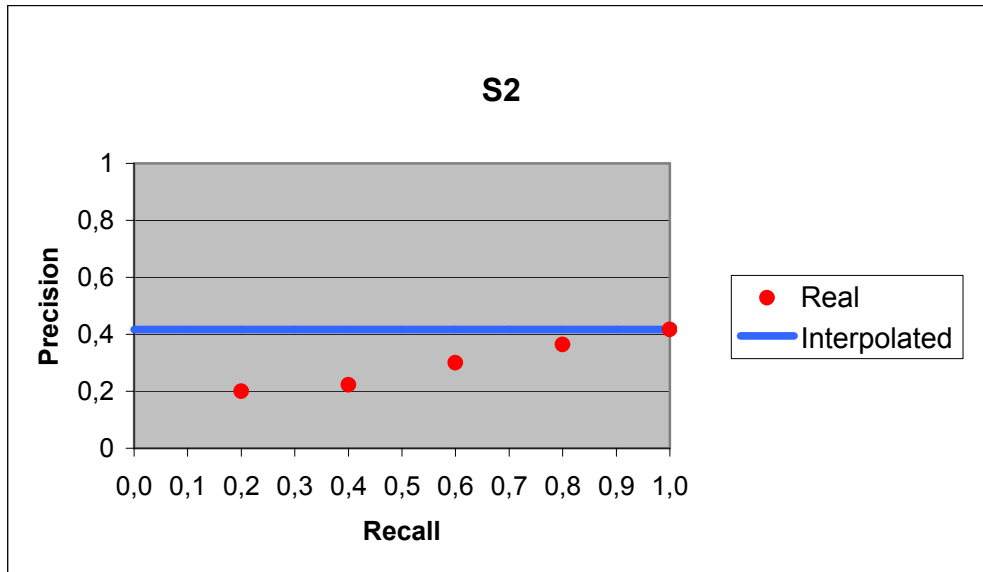
2° Συναφές: $Recall(S2) = \frac{2}{5} = 0.4$ $Precision(S2) = \frac{2}{9} = 0.222$

3° Συναφές: $Recall(S2) = \frac{3}{5} = 0.6$ $Precision(S2) = \frac{3}{10} = 0.3$

4° Συναφές: $Recall(S2) = \frac{4}{5} = 0.8$ $Precision(S2) = \frac{4}{11} = 0.364$

5° Συναφές: $Recall(S2) = \frac{5}{5} = 1$ $Precision(S2) = \frac{5}{12} = 0.417$

Και επομένως:



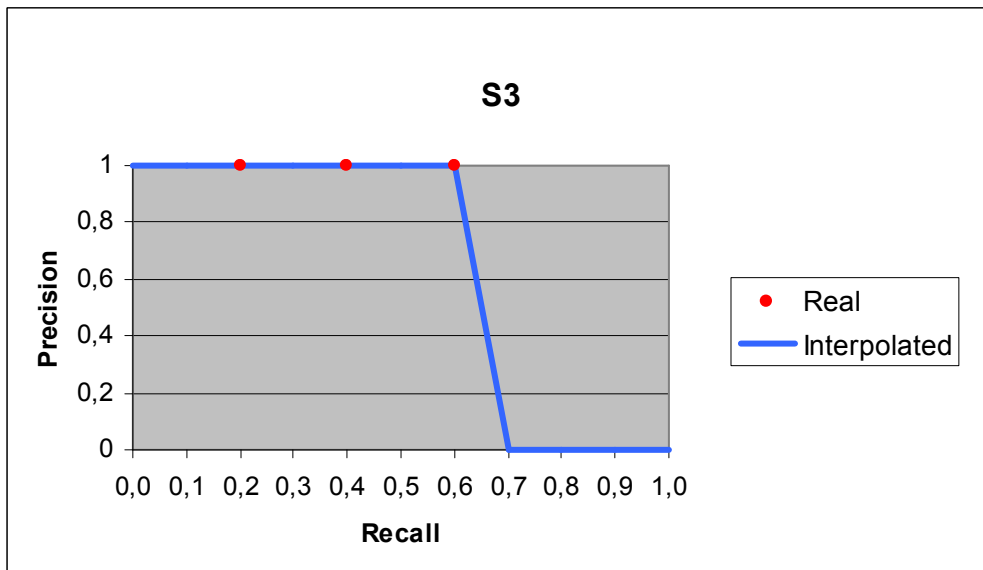
S3

Το S3 ανακτά συνολικά 3 έγγραφα

1° Συναφές: $Recall(S3) = \frac{1}{5} = 0.2$ $Precision(S3) = \frac{1}{1} = 1$

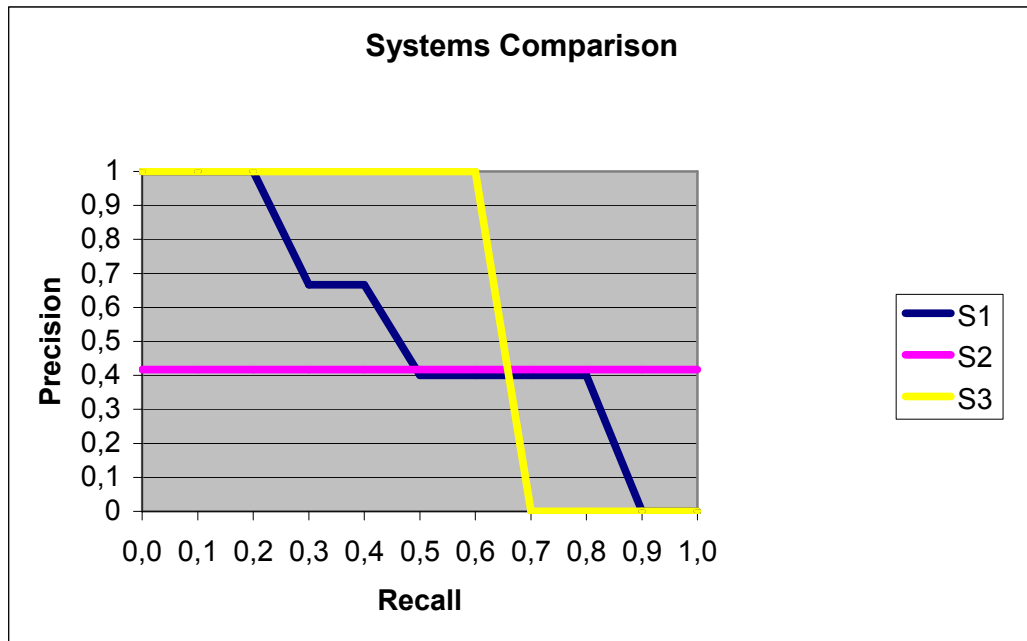
2° Συναφές: $Recall(S3) = \frac{2}{5} = 0.4$ $Precision(S3) = \frac{2}{2} = 1$

3° Συναφές: $Recall(S3) = \frac{3}{5} = 0.6$ $Precision(S3) = \frac{3}{3} = 1$



Εάν θέλουμε να διαλέξουμε ένα σύστημα κοιτώντας μόνο τις *P/R curves* τότε σίγουρα θα θέλαμε ένα γράφημα το οποίο θα έχει υψηλές τιμές ακρίβειας για υψηλές τιμές ανάκλησης. Επομένως ως καλύτερο σύστημα μπορούμε να επιλέξουμε αυτό του οποίου η καμπύλη προσεγγίζει αυτή την συμπεριφορά. Εναλλακτικά μπορούμε να επιλέξουμε αυτό του οποίου

το εμβαδό (η περιοχή κάτω και δεξιά από την καμπύλη) είναι μεγαλύτερο μιας και μεγαλύτερο εμβαδόν σημαίνει μεγαλύτερες τιμές ακρίβειας και ανάκλησης.



Επομένως φαίνεται ότι το προτιμότερο σύστημα είναι το S3.

Recall/Fallout Curves

S1

1° Συναφές: $Recall(S1) = \frac{1}{5} = 0.2$

$Fallout(S1) = \frac{0}{15} = 0$

2° Συναφές: $Recall(S1) = \frac{2}{5} = 0.4$

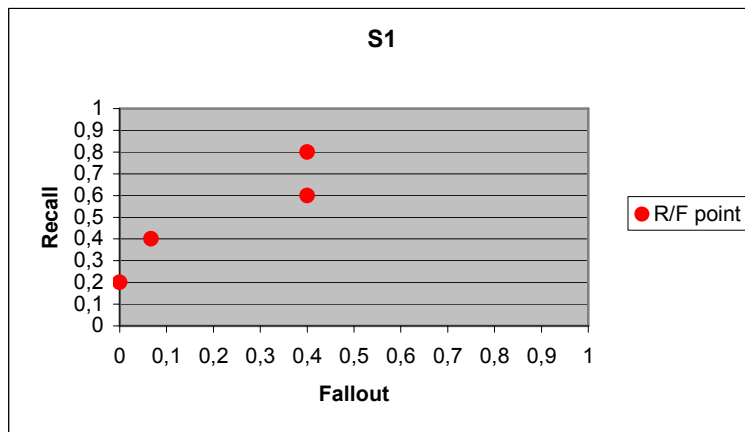
$Fallout(S1) = \frac{1}{15} = 0.067$

3° Συναφές: $Recall(S1) = \frac{3}{5} = 0.6$

$Fallout(S1) = \frac{6}{15} = 0.4$

4° Συναφές: $Recall(S1) = \frac{4}{5} = 0.8$

$Fallout(S1) = \frac{6}{15} = 0.4$



S2

1° Συναφές: $Recall(S2) = \frac{1}{5} = 0.2$

$$Fallout(S2) = \frac{4}{15} = 0.267$$

2° Συναφές: $Recall(S2) = \frac{2}{5} = 0.4$

$$Fallout(S2) = \frac{7}{15} = 0.467$$

3° Συναφές: $Recall(S2) = \frac{3}{5} = 0.6$

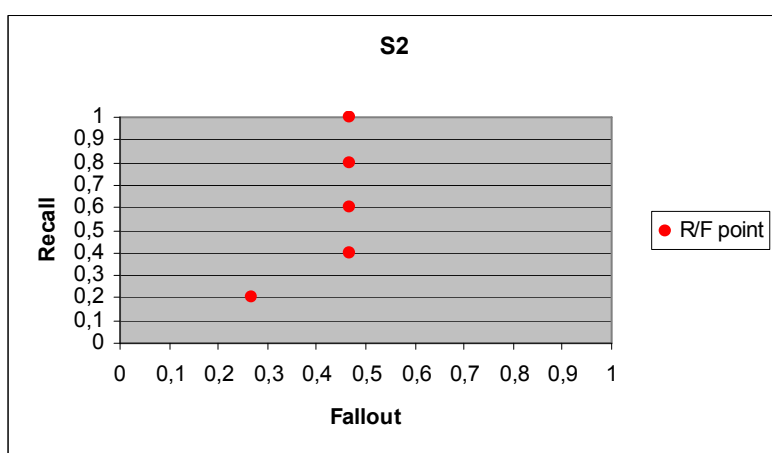
$$Fallout(S2) = \frac{7}{15} = 0.467$$

4° Συναφές: $Recall(S2) = \frac{4}{5} = 0.8$

$$Fallout(S2) = \frac{7}{15} = 0.467$$

5° Συναφές: $Recall(S2) = \frac{5}{5} = 1$

$$Fallout(S2) = \frac{7}{15} = 0.467$$



S3

1° Συναφές: $Recall(S3) = \frac{1}{5} = 0.2$

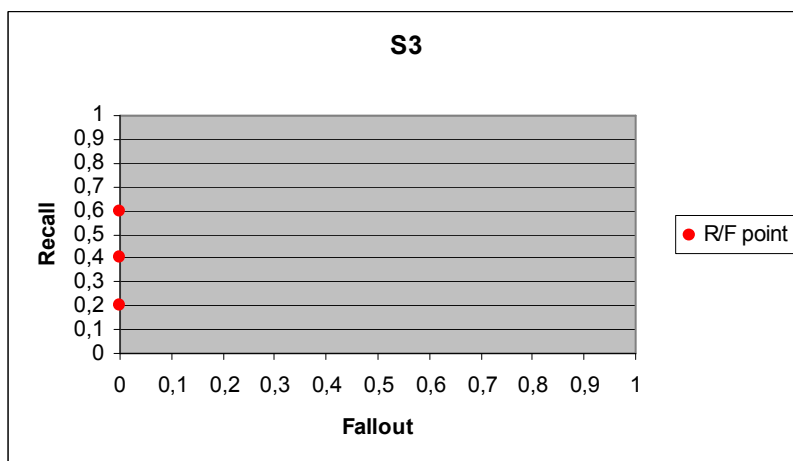
$$Fallout(S3) = \frac{0}{15} = 0$$

2° Συναφές: $Recall(S3) = \frac{2}{5} = 0.4$

$$Fallout(S3) = \frac{0}{15} = 0$$

3° Συναφές: $Recall(S3) = \frac{3}{5} = 0.6$

$$Fallout(S3) = \frac{0}{15} = 0$$



Ένα καλό σύστημα έχει χαμηλές τιμές *Fallout* (ιδανικά 0) για όλες τιμές του *Recall* ενώ το *Recall* πρέπει να προσεγγίζει το 1. Έτσι λοιπόν θεωρούμε ένα σύστημα καλό όταν οι τιμές που παίρνει στον άξονα των *X* (*Fallout*) είναι όσο το δυνατόν χαμηλότερες (και άρα ανακτά πολύ μικρό πλήθος από μη συναφή έγγραφα) όσο οι τιμές που παίρνει στον άξονα των *Y* (*Recall*) είναι όσο γίνεται μεγαλύτερες (και άρα ανακτά μεγάλο πλήθος από συναφών εγγράφων). Για παράδειγμα ένα ιδανικό σύστημα έχει *Recall*=1 & *Fallout*=0. Έτσι λοιπόν κοιτώντας τις καμπύλες *Recall/Fallout* μπορούμε εύκολα να καταλήξουμε στο ίδιο συμπέρασμα που καταλήξαμε κοιτώντας τις *Recall/Fallout* καμπύλες, ότι δηλαδή το σύστημα *S3* έχει καλύτερη συμπεριφορά σε σχέση με τα άλλα 2 επειδή επιστρέφει τα συναφή έγγραφα πρώτα.

Άσκηση 3

Θεωρήστε ότι έχουμε μια συλλογή αξιολόγησης που αποτελείται από ένα σύνολο εγγράφων *V* και ένα σύνολο επερωτήσεων *Q*. Για κάθε ζεύγος $q, u_i \in Q \times V$ έχει οριστεί το $label_q(u_i)$, το οποίο εκφράζει τη συνάφεια του u_i ως προς το q σε μία κλίμακα από το 0 έως το 5 (0,1,2,3,4,5) όπου 0=χειρότερο, ..., 5=καλύτερο. Έστω ότι θέλουμε να αξιολογήσουμε την απόκριση ενός συστήματος σε μία επερώτηση q . Έστω ότι τα κορυφαία k στοιχεία της απόκρισης είναι το διατεταγμένο σύνολο $\langle u_1, \dots, u_k \rangle$. Για την αξιολόγηση του έχουμε στην διάθεση μας τα εξής μέτρα:

(α)

$$M_1(k) = \sum_{i=1}^k (2^{label_q(u_i)} - 1)$$

(β)

$$M_2(k) = \sum_{i=1}^k \frac{1}{\log_2(i+1)} (label_q(u_i) - 1)$$

(γ)

$$M_2(k) = \sum_{i=1}^k \frac{1}{\log_2(i+1)} (2^{label_q(u_i)} - 1)$$

Σχολιάστε σύντομα το κάθε ένα. Συγκρίνετε τα και δικαιολογείστε ποιο θα χρησιμοποιούσατε. Στηρίξτε τις απόψεις δίνοντας μικρά παραδείγματα.

Λύση

(α) Παρατηρούμε ότι για κάθε έγγραφο η τιμή που θα προστίθεται στο συνολικό άθροισμα θα είναι {0, 1, 3, 7, 15, 31}. Επομένως εάν η απάντηση έχει μη συναφή έγγραφα τότε αυτά δεν θα συνεισφέρουν καθόλου στο γενικό άθροισμα ενώ εάν έγγραφο είναι πολύ καλό (δηλαδή $label_q(u_i) = 5$) συνεισφέρει πολύ περισσότερο από άλλα λιγότερο σχετικά στο γενικό άθροισμα. Επομένως βλέπουμε ότι αυτό το μέτρο πριμοδοτεί ιδιαίτερα τα πολύ συναφή έγγραφα κάτι το οποίο είναι γενικά θεμιτό. Όμως δεν λαμβάνει υπ' όψιν του καθόλου την διάταξη των εγγράφων. Συγκεκριμένα το ίδιο θα συνεισφέρει στο γενικό άθροισμα ένα πολύ καλό έγγραφο είτε εμφανίζεται στην πρώτη είτε εμφανίζεται στην τελευταία θέση του διατεταγμένου συνόλου.

Έτσι για παράδειγμα έστω ότι έχουμε τις διατεταγμένες απαντήσεις

$ans_1 = \langle 5, 5, 0, 0, 0 \rangle$ και $ans_2 = \langle 0, 0, 0, 5, 5 \rangle$ (όπου αντί για έγγραφα έχουν τοποθετηθεί τα αντίστοιχα $label_q(u_i)$). Τόσο η ans_1 όσο και η ans_2 έχουν 2 πολύ καλά έγγραφα και 3 κακά. Όμως η πρώτη απάντηση είναι σαφώς καλύτερη από την δεύτερη αφού επιστρέφει τα καλά έγγραφα πρώτα σε αντίθεση με τη άλλη που τα επιστρέφει στο τέλος. Κι όμως το μέτρο M_1 μας δίνει

- $M_1(5) = (31 + 31 + 0 + 0 + 0) = 62$ για την ans_1
- $M_1(5) = (0 + 0 + 0 + 31 + 31) = 62$ για την ans_2

(β) Το δεύτερο μέτρο που έχουμε στην διάθεση μας παρατηρούμε ότι πλέον λαμβάνει υπ' όψιν την θέση στην οποία εμφανίζεται το κάθε έγγραφο. Συγκεκριμένα η βαθμολογία του κάθε εγγράφου διαιρείται με ένα συντελεστή $\log_2(i+1)$. Επομένως το πρώτο έγγραφο θα διαιρεθεί με $\log_2(2)$, το δεύτερο με $\log_2(3)$ κ.ο.κ. Έτσι πλέον ένα έγγραφο που εμφανίζεται στις πρώτες θέσεις της διάταξης συνεισφέρει περισσότερο από ένα έγγραφο που βρίσκεται σε κάποια θέση πιο μετά. Ωστόσο βλέπουμε ότι πλέον τα έγγραφα ανάλογα με την βαθμολογία τους παίρνουν τιμές : $\{-1, 0, 1, 2, 3, 4\}$. Έτσι λοιπόν τα πολύ καλά έγγραφα δεν πριμοδοτούντε λίγο παραπάνω (όπως με το μέτρο M_1), Για παράδειγμα έστω οι απαντήσεις:

$ans_1 = \langle 2, 2, 3, 2, 2 \rangle$ και $ans_2 = \langle 1, 1, 1, 5, 5 \rangle$. Αν χρησιμοποιήσουμε το μέτρο M_2 θα βρούμε ότι

- $M_2(5) = (1 + 0.63 + 1 + 0.43 + 0.39) = 3.45$ για την ans_1
- $M_2(5) = (0 + 0 + 0 + 1.72 + 1.55) = 3.27$ για την ans_2

Παρατηρούμε πόσο κοντά είναι οι τιμές του μέτρου M_2 για τις 2 απαντήσεις και μάλιστα θεωρεί την απάντηση ans_1 καλύτερη (η οποία το πιο συναφές έγγραφο που περιέχει είναι ένα βαθμού 3 στην 3^η θέση της διάταξης) από την ans_2 (η οποία περιέχει 2 έγγραφα βαθμού 5 στις θέσεις 4, 5).

(γ) Το τρίτο μέτρο λαμβάνει υπ' όψιν την θέση στην οποία βρίσκεται ένα έγγραφο αφού διαιρεί με ένα συντελεστή $\log_2(i+1)$ (όπως και το μέτρο M_2). Επίσης κάθε έγγραφο συνεισφέρει εκθετικά με τιμές $\{0, 1, 3, 7, 15, 31\}$ (όπως και το μέτρο M_1). Έτσι τα πολύ συναφή έγγραφα συνεισφέρουν περισσότερο στο άθροισμα απ' ότι τα λιγότερο συναφή. Τέλος τα μη συναφή έγγραφα δεν "καταστρέφουν" το γενικό άθροισμα μειώνοντας το. Επομένως το M_3 φαίνεται να συνδυάζει τα προτερήματα των 2 παραπάνω. Έστω για παράδειγμα οι απαντήσεις $ans_1 = \langle 2, 2, 3, 2, 2 \rangle$ και $ans_2 = \langle 1, 1, 1, 5, 5 \rangle$

Χρησιμοποιώντας το μέτρο M_3 έχουμε:

- $M_3(5) = (3 + 1.89 + 3.5 + 1.29 + 1.16) = 10.84$ για την ans_1
- $M_3(5) = (1 + 0.63 + 0.5 + 13.35 + 11.99) = 27.47$ για την ans_2

Επομένως το μέτρο M_1 είναι καλύτερο από πλευράς συνάφειας αλλά δεν λαμβάνει υπ' όψιν την θέση που βρίσκεται ένα έγγραφο ενώ το μέτρο M_2 λαμβάνει υπ' όψιν του τις θέσεις στις οποίες βρίσκονται τα έγγραφα αλλά δεν είναι τόσο καλό από πλευράς συνάφειας. Τέλος το M_3 συνδυάζει τα πλεονεκτήματα και των 2 παραπάνω και έτσι είναι η καταλληλότερη επιλογή.

Άσκηση 4

Έστω ότι η συλλογή αξιολόγησης αποτελείται από 100 έγγραφα $\{d_1, \dots, d_{100}\}$ και γνωρίζουμε ότι υπάρχουν 3 έγγραφα της συλλογής, συγκεκριμένα τα $\{d_1, d_2, d_3\}$, που είναι συναφή με την επερώτηση q . Θέλουμε να αξιολογήσουμε την αποτελεσματικότητα τριών συστημάτων $S1$, $S2$ και $S3$ τα οποία επιστρέφουν ως απάντηση έγγραφα συνοδευμένα από ένα βαθμό συνάφειας. Υποβάλλουμε σε κάθε σύστημα την επερώτηση q και λαμβάνουμε τις εξής απαντήσεις:

$$Ans(S1, q) = \langle d_1, \{d_2, d_{20} - d_{40}\}, d_3 \rangle$$

$$Ans(S2, q) = \langle d_1, d_2, d_3 \rangle$$

$$Ans(S3, q) = \langle \{d_1, d_8\}, d_2, d_3 \rangle$$

Η απάντηση $\langle \{d_1, d_8\}, d_2, d_3 \rangle$ σημαίνει ότι τα d_1, d_8 ισοβαθούν στην πρώτη θέση (άρα έλαβαν το μεγαλύτερο βαθμό συνάφειας). Η απάντηση $\langle d_1, \{d_2, d_{20} - d_{40}\}, d_3 \rangle$ σημαίνει ότι το d_1 έλαβε το μεγαλύτερο βαθμό, ενώ μετά ακολουθεί μια ομάδα από 22 έγγραφα τα οποία ισοβαθούν και στο τέλος της κατάταξης βρίσκεται το d_3 . Για κάθε ένα από τα 3 συστήματα απαντήστε στα ακόλουθα ερωτήματα: (α) Ποια είναι η R-Ακρίβεια (R-Precision); (β) Ποιο είναι το αναμενόμενο μήκος αναζήτησης για να βρούμε 2 συναφή;

Λύση

(α) Η R-ακρίβεια είναι η ακρίβεια στην R θέση της διάταξης, όπου R είναι το πλήθος των συναφών εγγράφων. Εδώ έχουμε $R=3$ επομένως

S1

Στην 3^η θέση της διάταξης των αποτελεσμάτων παρατηρούμε ότι μπορεί να υπάρχουν 22 διαφορετικά έγγραφα ($\{d_2, d_{20} - d_{40}\}$). Επιπλέον παρατηρούμε ότι μέσα σε αυτά τα 22 έγγραφα υπάρχει ένα (το d_2) το οποίο είναι συναφές με την επερώτηση q . Επομένως οι διαφορετικές θέσεις στις οποίες μπορεί να βρίσκεται το d_2 μπορεί να μας δώσουν διαφορετικές τιμές για την R-ακρίβεια του συστήματος $S1$. Συγκεκριμένα:

- $R - Precision(S1) = \frac{2}{3} = 0.667$ όταν το d_2 βρίσκεται στην 2^η ή στην 3^η θέση
- $R - Precision(S1) = \frac{1}{3} = 0.333$ όταν το d_2 βρίσκεται σε κάποια από τις θέσεις $\{4^η, 5^η, \dots, 23^η\}$

Συνολικά οι διαφορετικές θέσεις στις οποίες μπορεί να βρίσκεται το d_2 είναι 22 και

- Σε 2 από αυτές τις θέσεις $R - Precision(S1) = \frac{2}{3} = 0.667$
- Σε 20 από αυτές τις θέσεις $R - Precision(S1) = \frac{1}{3} = 0.333$

$$\text{Επομένως συνολικά } R - Precision(S1) = \frac{2}{22} \cdot \frac{2}{3} + \frac{20}{22} \cdot \frac{1}{3} = \frac{4}{66} + \frac{20}{66} = \frac{24}{66} = 0.364$$

S2

Εδώ παρατηρούμε ότι στην 3^η θέση της διάταξης συναντάμε το 3^ο συναφές έγγραφο επομένως:

$$R - Precision(S2) = \frac{3}{3} = 1$$

S3

Στην 3^η θέση της διάταξης συναντάμε το έγγραφο d_2 . Μέχρι εκείνη τη θέση έχουμε ήδη δει ένα συναφές έγγραφο (d_1) και ένα μη συναφές (d_8). Επομένως:

$$R - Precision(S3) = \frac{2}{3} = 0.667$$

(β) Το αναμενόμενο μήκος αναζήτησης είναι το πλήθος των μη συναφών εγγράφων που πρέπει να αναζητήσουμε προκειμένου να έχουμε ένα συγκεκριμένο πλήθος συναφών εγγράφων. Θέλουμε το μήκος αναζήτησης για να έχουμε 2 συναφή έγγραφα.

S1

Στην απάντηση του συστήματος $S1$ το 2^ο συναφές έγγραφο (d_2) παρατηρούμε ότι μπορεί να βρίσκεται σε κάποια από τις θέσεις $\{2^1, 3^1, \dots, 23^1\}$. Επομένως ανάλογα με την θέση που θα βρίσκεται το δεύτερο συναφές έγγραφο μπορεί να έχουμε διαφορετικό μήκος αναζήτησης. Συγκεκριμένα:

- Εάν το d_2 βρίσκεται στην 2^η θέση τότε το $SearchLength(S1) = 0$
- Εάν το d_2 βρίσκεται στην 3^η θέση τότε το $SearchLength(S1) = 1$
- ...
- Εάν το d_2 βρίσκεται στην 23^η θέση τότε το $SearchLength(S1) = 21$

$$\text{Επομένως } SearchLength(S1) = \frac{0+1+\dots+21}{22} = 10.5$$

S2

Το δεύτερο συναφές έγγραφο το συναντάμε στην 2^η θέση και επομένως $SearchLength(S2) = 0$

S3

Το δεύτερο συναφές έγγραφο το συναντάμε στην 3^η θέση και επομένως έχουμε ήδη δει ένα μη συναφές. Επομένως $SearchLength(S3) = 1$