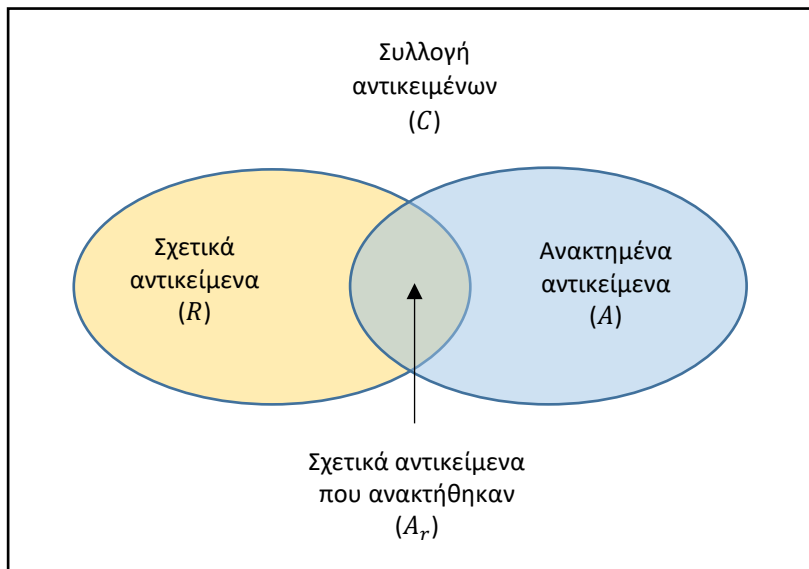


Αποτίμηση Αποτελεσματικότητας

Γενικά

Έστω C μια συλλογή αντικειμένων (objects) π.χ. αντικείμενα σε μια εικόνα ή έγγραφα ή λέξεις σε μια σελίδα. Με βάση κάποιο ερώτημα/κριτήριο αναζητούνται κάποια σχετικά αντικείμενα R (related objects) που είναι υποσύνολο του C . Για παράδειγμα, κάποια συγκεκριμένα έγγραφα σε μια συλλογή εγγράφων ή κάποια λέξη στο σύνολο των λέξεων μιας σελίδας. Έστω ότι κάποια μέθοδος επιστρέφει ως απάντηση στο ερώτημα/κριτήριο αναζήτησης ένα σύνολο ανακτημένων αντικειμένων A (detected objects) που αποτελεί επίσης υποσύνολο του C . Με A_r συμβολίζουμε το υποσύνολο του A που περιέχει τα σχετικά ως προς το ερώτημα ανακτημένα αντικείμενα. Στο παρακάτω σχήμα απεικονίζονται σχηματικά οι διάφορες κατηγορίες.



Σχήμα 1 Διάγραμμα Venn για αναπαράσταση της συλλογής και των διαφόρων κατηγοριών αντικειμένων.

Πίνακας σύγχυσης

Η απόδοση της μεθόδου εντοπισμού μπορεί να εκφρασθεί ποσοτικά κατηγοριοποιώντας τα αποτελέσματα στις παρακάτω τέσσερις κατηγορίες:

Πίνακας 1 Κατηγορίες συσχέτισης σχετικών και ανακτημένων αντικειμένων

TP , true positives	A_r	Σχετικά αντικείμενα που ανακτήθηκαν
FP , false positives	$A - A_r$	Μη σχετικά αντικείμενα που ανακτήθηκαν
FN , false negatives	$R - A_r$	Σχετικά αντικείμενα που δεν ανακτήθηκαν
TN , true negatives	$C - (R \cup A)$	Μη σχετικά αντικείμενα που δεν ανακτήθηκαν

Με βάση τα TP , FN , FP και TN σχηματίζεται ο πίνακας σύγχυσης (**confusion matrix**):

		Σχετικά αντικείμενα (ground truth)	
		Σωστό	Λάθος
Ανακτημένα αντικείμενα	Σωστό	TP	FP
	Λάθος	FN	TN

Σχήμα 2 Πίνακας σύγχυσης

Ανάκληση και Ακρίβεια

Δύο μετρικές που χρησιμοποιούνται συχνά κατά την διαδικασία αποτίμησης είναι η ακρίβεια και η ανάκληση. Οι μετρικές αυτές μπορούν να υπολογιστούν με βάση τα δεδομένα του confusion matrix.

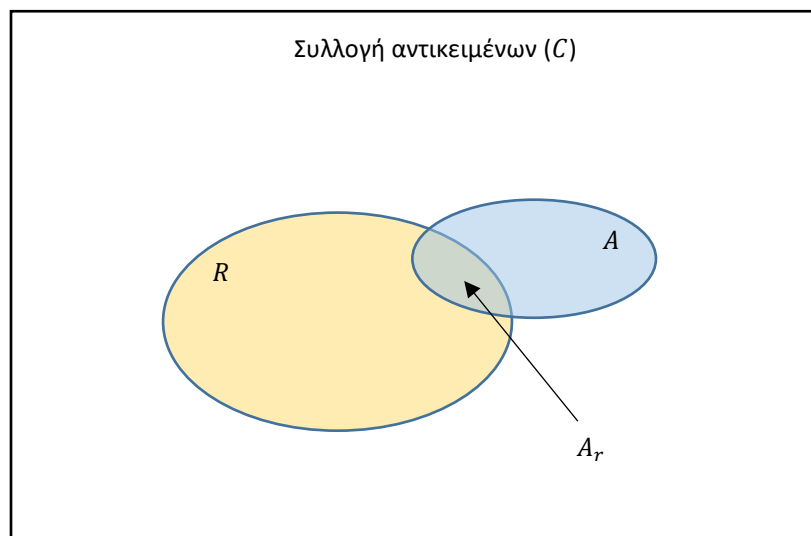
Η ανάκληση (**recall**) εκφράζεται από τον λόγο του πλήθους των σχετικών ανακτημένων αντικειμένων προς το πλήθος των σχετικών αντικειμένων στην συλλογή:

$$Recall = \frac{\text{αριθμός σχετικών αντικειμένων που ανακτήθηκαν}}{\text{αριθμός σχετικών αντικειμένων στην συλλογή}} = \frac{|A_r|}{|R|} = \frac{TP}{TP + FN}$$

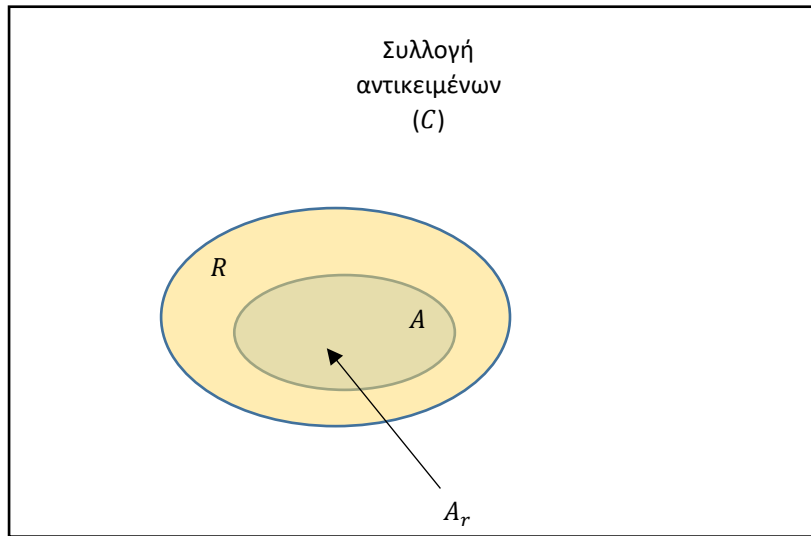
Η ακρίβεια (**precision**) εκφράζεται από τον λόγο του πλήθους των σχετικών ανακτημένων αντικειμένων προς το πλήθος των ανακτημένων αντικειμένων:

$$Precision = \frac{\text{αριθμός σχετικών αντικειμένων που ανακτήθηκαν}}{\text{αριθμός ανακτημένων αντικειμένων}} = \frac{|A_r|}{|A|} = \frac{TP}{TP + FP}$$

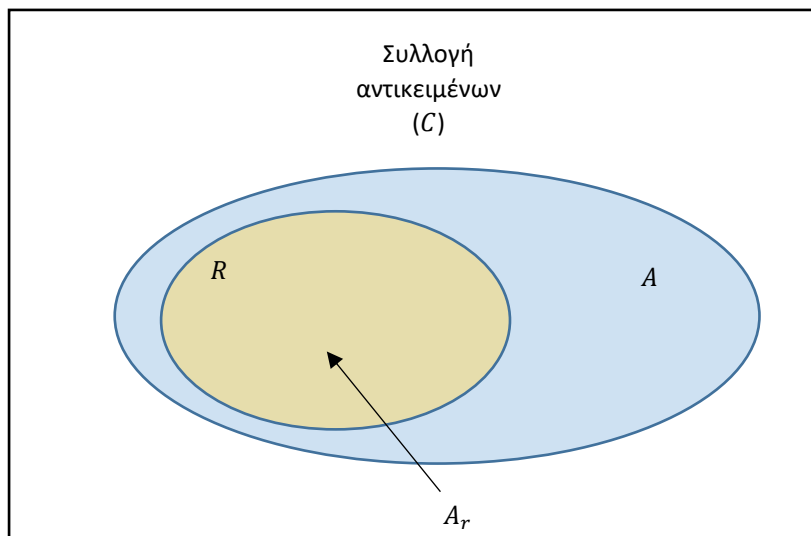
Στα παρακάτω σχήματα απεικονίζονται τέσσερις περιπτώσεις ανάκλησης και ακρίβειας.



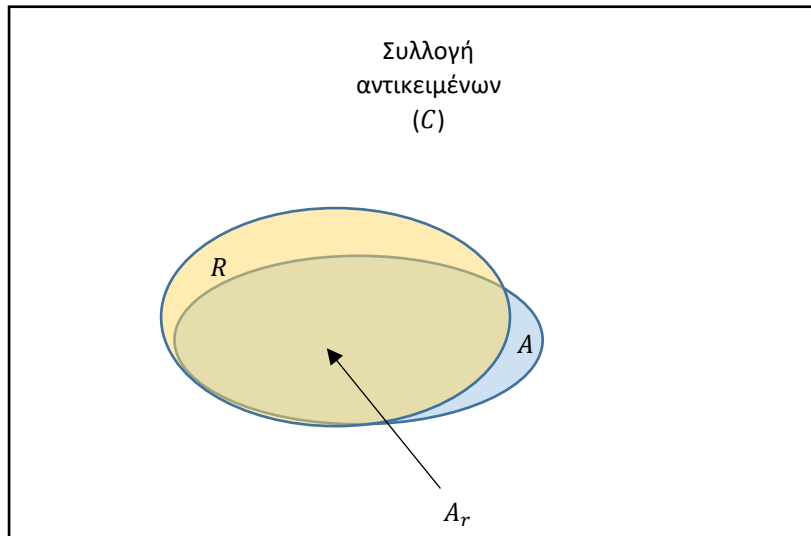
Σχήμα 3 Χαμηλή ανάκληση, χαμηλή ακρίβεια. Τα ανακτημένα αντικείμενα περιέχουν λίγα από τα σχετικά αντικείμενα.



Σχήμα 4 Χαμηλή ανάκληση, υψηλή ακρίβεια. Τα ανακτημένα αντικείμενα είναι όλα σχετικά αλλά υπάρχουν και άλλα σχετικά αντικείμενα που δεν έχουν ανακτηθεί.



Σχήμα 5 Υψηλή ανάκληση, χαμηλή ακρίβεια. Τα σχετικά αντικείμενα έχουν ανακτηθεί όλα αλλά ταυτόχρονα έχουν ανακτηθεί και μη σχετικά αντικείμενα.



Σχήμα 6 Υψηλή ανάκληση, υψηλή ακρίβεια. Τα σχετικά αντικείμενα έχουν ανακτηθεί σχεδόν όλα μαζί με λίγα μη σχετικά.

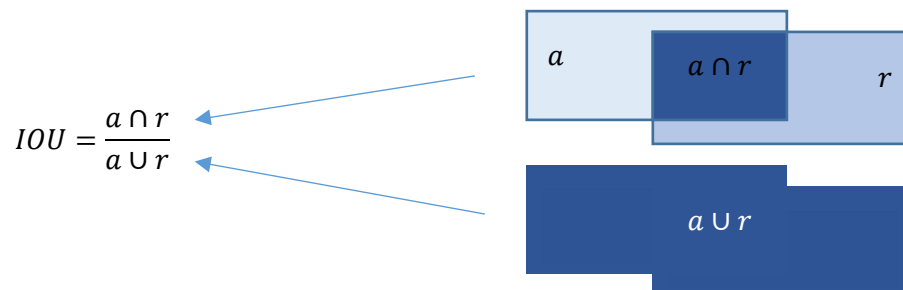
F-Measure

Μια μετρική που συνδυάζει την ανάκληση και την ακρίβεια σε μία ποσότητα είναι ο αρμονικός μέσος τους (**F-Measure** ή **F1 score**) και ορίζεται ως

$$F_1 = 2 \cdot \frac{1}{\frac{1}{\text{Recall}} + \frac{1}{\text{Precision}}} = 2 \cdot \frac{\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}}$$

Intersection Over Union

Συχνά σε προβλήματα ανάκτησης σε εικόνες, τα αντικείμενα προσδιορίζονται μέσω ενός περιγεγραμμένου ορθογωνίου (**bounding box**), π.χ αντικείμενα σε μια φυσική εικόνα, λέξεις σε μια εικόνα εγγράφου κλπ. Για να προσδιοριστεί κατά πόσο το ανακτημένο αντικείμενο $a \in A$ αντιστοιχεί σε κάποιο σχετικό αντικείμενο $r \in R$ χρησιμοποιείται η μετρική **Intersection Over Union (IOU)** που βασίζεται στα αντίστοιχα δύο ορθογώνια και ορίζεται ως



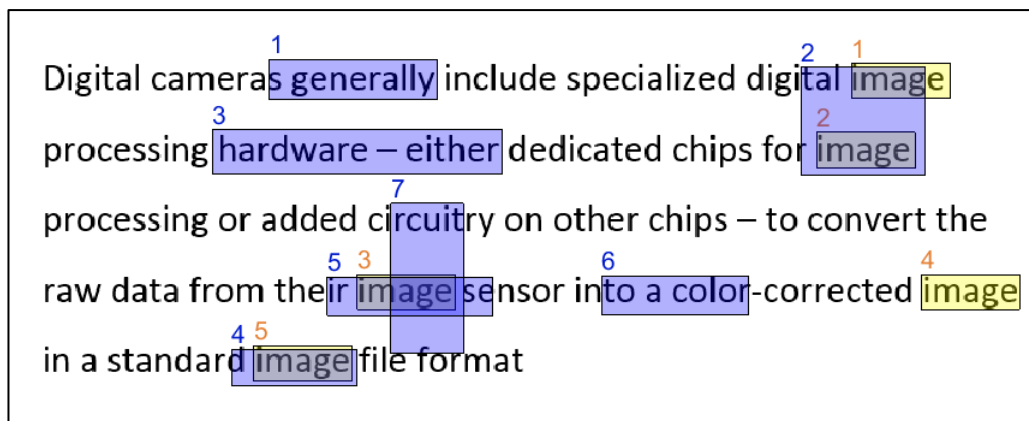
Σχήμα 7 Μετρική Intersection Over Union.

Η τιμή του IOU κυμαίνεται μεταξύ 0 (καθόλου επικάλυψη) και 1 (πλήρης ταύτιση). Εάν $IOU \geq T$, δηλαδή εάν υπερβαίνει κάποιο κατώφλι T (παράμετρος του συστήματος που

συνήθως ορίζεται από τον χρήστη) τότε θεωρείται ότι το σχετικό αντικείμενο r έχει εντοπισθεί σωστά από το ανακτημένο αντικείμενο a .

Παράδειγμα

Στο παρακάτω παράδειγμα αναζητούμε την λέξη «image» σε μια εικόνα κειμένου με συνολικά $N_C = 40$ λέξεις. Με βάση κάποιο ερώτημα/κριτήριο, οι αναζητούμενες σχετικές λέξεις (related words) είναι $N_R = 5$ και είναι μαρκαρισμένες με ορθογώνιο κίτρινου χρώματος. Έστω ότι η εφαρμοζόμενη μέθοδος ανάκτησε $N_A = 7$ λέξεις (detected words) που απεικονίζονται με ορθογώνιο μπλε χρώματος.



Σχήμα 8 Παράδειγμα εικόνας κειμένου. Με κίτρινο χρώμα απεικονίζονται οι αναζητούμενες σχετικές λέξεις και ενώ με μπλε χρώμα απεικονίζονται οι ανακτημένες λέξεις.

Ο υπολογισμός του IOU για κάθε ζεύγος σχετικής και ανακτημένης λέξης δίνει έναν πίνακα $N_A \times N_R = 7 \times 5$ τιμών IOU :

Πίνακας 2 IOU για όλα τα ζεύγη σχετικών και αναζητούμενων λέξεων

		Σχετικές λέξεις				
		1	2	3	4	5
Ανακτημένες λέξεις	1	0	0	0	0	0
	2	0.17	0.28	0	0	0
	3	0	0	0	0	0
	4	0	0	0	0	0.68
	5	0	0	0.51	0	0
	6	0	0	0	0	0
	7	0	0	0.2	0	0

1^η περίπτωση

Εάν ορίσουμε ως κατώφλι επικάλυψης $T = 0.5$ τότε ο παραπάνω πίνακας γίνεται:

Πίνακας 3 Αποτελέσματα ανάκτησης για κατώφλι επικάλυψης $T = 0.5$. Τα TP , FP και FN απεικονίζονται με πράσινο, κόκκινο και μωβ χρώμα, αντίστοιχα.

		Σχετικές λέξεις					
		1	2	3	4	5	
Ανακτημένες λέξεις	1	0	0	0	0	0	0
	2	0	0	0	0	0	0
	3	0	0	0	0	0	0
	4	0	0	0	0	1	1
	5	0	0	1	0	0	1
	6	0	0	0	0	0	0
	7	0	0	0	0	0	0

↓

0	0	1	0	1
---	---	---	---	---

Οπότε,

True positives $TP = 2$, False positives $FP = 5$, False negatives $FN = 3$ και True negatives $TN = 30$ (οι υπόλοιπες λέξεις του κειμένου).

Ο πίνακας σύγχυσης γίνεται

		Σχετικά αντικείμενα (ground truth)	
		Σωστό	Λάθος
Ανακτημένα αντικείμενα	Σωστό	2	5
	Λάθος	3	30

Η ανάκληση και ακρίβεια είναι

$$Recall = \frac{TP}{TP + FN} = \frac{2}{2 + 3} = \frac{2}{5} = 0.40$$

$$Precision = \frac{TP}{TP + FP} = \frac{2}{2 + 5} = \frac{2}{7} = 0.29$$

Ο αρμονικός μέσος γίνεται

$$F_1 = 2 \cdot \frac{Recall \cdot Precision}{Recall + Precision} = 2 \cdot \frac{0.4 \cdot 0.29}{0.4 + 0.29} = 0.34$$

2^η περίπτωση

Εάν ορίσουμε ως κατώφλι επικάλυψης $T = 0.2$ και το εφαρμόσουμε στον Πίνακα 2 τότε τα αποτελέσματα της ανάκτησης είναι:

Πίνακας 4 Αποτελέσματα ανάκτησης για για κατώφλι επικάλυψης $T = 0.2$

		Σχετικές λέξεις					
		1	2	3	4	5	
Ανακτημένες λέξεις	1	0	0	0	0	0	0
	2	0	1	0	0	0	1
	3	0	0	0	0	0	0
	4	0	0	0	0	0	1
	5	0	0	1	0	0	1
	6	0	0	0	0	0	0
	7	0	0	1	0	0	0

↓

0	1	1	0	1
---	---	---	---	---

Οπότε,

True positives $TP = 3$, False positives $FP = 4$, False negatives $FN = 2$ και True negatives $TN = 31$ (οι υπόλοιπες λέξεις του κειμένου).

Ο πίνακας σύγκρισης γίνεται

		Σχετικά αντικείμενα (ground truth)	
		Σωστό	Λάθος
Ανακτημένα αντικείμενα	Σωστό	3	4
	Λάθος	2	31

Η ανάκληση και ακρίβεια είναι

$$Recall = \frac{TP}{TP + FN} = \frac{3}{3 + 2} = \frac{3}{5} = 0.60$$

$$Precision = \frac{TP}{TP + FP} = \frac{3}{3 + 4} = \frac{3}{7} = 0.43$$

Ο αρμονικός μέσος γίνεται

$$F_1 = 2 \cdot \frac{Recall \cdot Precision}{Recall + Precision} = 2 \cdot \frac{0.6 \cdot 0.43}{0.6 + 0.43} = 0.50$$