

2^η Σειρά Ασκήσεων

Μοντέλα Ανάκτησης και Αξιολόγηση Ανάκτησης
Ημερομηνία Παράδοσης: Δευτέρα 6 Απριλίου 2009

Άσκηση 1

(Συνολικά 2 ομάδες από 3 άτομα η κάθε μια) Διαβάστε το μοντέλο *MAP* που παρέχει μια *θεωρητική εκτίμηση* των ποσοτήτων p_i και u_i (σελ. 207-208) καθώς και το *Okapi BM25* (σελ. 213-215) από το βιβλίο των Manning et. al. Χρησιμοποιείτε το παράδειγμα του προηγούμενου μαθήματος (τα 5 έγγραφα και τις 4 ερωτήσεις στο τέλος των διαφανειών). Κάθε ομάδα θα ετοιμάσει μια σύντομη παρουσίαση για ένα από τα δύο μοντέλα.

Άσκηση 2

(Ομάδες των 2 ατόμων) 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 4, συγκεκριμένα τα $\{d_2, d_6, d_7, d_{11}\}$. Θέλουμε να αξιολογήσουμε την αποτελεσματικότητα τριών συστημάτων S_1, S_2 και S_3 .

Για το λόγο αυτό υποβάλλουμε σε κάθε σύστημα την επερώτηση q και λαμβάνουμε τις εξής απαντήσεις :

$Ans(S_1, q) = \langle d_1, d_2, d_7, d_{11}, d_5, d_{10}, d_{12}, d_{14}, d_3, d_4 \rangle$

$Ans(S_2, q) = \langle d_6, d_9, d_{10}, d_1, d_8, d_{11}, d_{12}, d_2, d_{17}, d_{15} \rangle$

$Ans(S_3, q) = \langle d_1, d_{11}, d_7, d_8, d_{15} \rangle$

Το αριστερότερο στοιχείο της κάθε απάντησης παριστάνει το υψηλότερα διαβαθμισμένο έγγραφο, δηλαδή αυτό που το σύστημα υπολόγισε ως το πιο συναφές με την επερώτηση q . Συγκρίνετε τα τρία αυτά συστήματα ως προς τα εξής μέτρα:

(α) Precision (Ακρίβεια)

(β) Recall (Ανάκληση)

(γ) F-Measure ($\beta = 2$ και $\beta = 0.5$)

(δ) R-Precision (R-Ακρίβεια)

(ε) Fallout

(στ) Δείξτε τις καμπύλες ακρίβειας -ανάκλησης σε 11 κανονικοποιημένα επίπεδα ανάκλησης.

(ζ) Δείξτε τις καμπύλες fallout-ανάκλησης σε 11 κανονικοποιημένα επίπεδα ανάκλησης.

(η) Αναμενόμενο μήκος αναζήτησης

Άσκηση 3

(Ομάδες των 2 ατόμων) Θεωρείστε τη *μέση ακρίβεια* στα σχετικά έγγραφα της απάντησης (average precision at the relevant document seen) που ορίζεται ως το άθροισμα των τιμών της ακρίβειας κάθε φορά που αλλάζει η ανάκληση (δηλαδή, κάθε φορά που βλέπουμε στην απάντηση ένα έγγραφο που είναι σχετικό) χωρίς να κάνουμε παρεμβολή - όπως ορίζεται στη σελίδα 52 των διαφανειών.

Θεωρείστε ότι έχετε ένα σύστημα αξιολόγησης που υπολογίζει τη μέση ακρίβεια κοιτάζοντας μόνο τα πρώτα 1.000 έγγραφα μιας διατεταγμένης λίστας. Οποιοδήποτε έγγραφο που δεν ανήκει στα 1.000 πρώτα θεωρείται ότι έχει διαβάθμιση (rank) άπειρο (infinity) δηλαδή η ακρίβεια για αυτό το σχετικό έγγραφο είναι 0 (δεν το θεωρούμε στην απάντηση). Θεωρείστε ότι τρέχετε ένα ερώτημα σε μια συλλογή ενός εκατομμυρίου (1.000.000) εγγράφων και ότι από αυτά τα σχετικά έγγραφα είναι 10 και μόνο 6 βρίσκονται στα 1.000 πρώτα της λίστας. Για αυτήν την ερώτηση ποια είναι η μεγαλύτερη και ποια η μικρότερη τιμή για τη μέση ακρίβεια.