**Άσκηση 1**

**Ακρίβεια και Ανάκληση**

Θεωρείστε μια συλλογή αξιολόγησης που αποτελείται από 50 τεκμήρια {d1,…,d50}.

Η συλλογή αξιολόγησης περιλαμβάνει μια επερώτηση q για την οποία γνωρίζουμε ότι τα τεκμήρια της συλλογής που είναι σχετικά με αυτήν είναι 10, συγκεκριμένα τα:   
{d1,d2,d6,d7,d11,d16,d19,d22,d25,d39}

Θέλουμε να αξιολογήσουμε την αποτελεσματικότητα δύο ΣΑΠ: S1 και S2.  
Για το λόγο αυτό υποβάλλουμε σε κάθε σύστημα την επερώτηση q και λαμβάνουμε τις εξής απαντήσεις :

Ans(S1,q)=<d1,d20,d7,d11,d5,d10,d12,d14,d2,d4,d39,d25,d16,d6,d41>  
Ans(S2,q)=<d6,d7,d10,d1,d8,d11,d18,d19,d2,d17,d15,d42,d22,d39,d16>  
Το αριστερότερο στοιχείο της κάθε απάντησης παριστάνει το υψηλότερα διαβαθμισμένο τεκμήριο, αυτό που το σύστημα υπολόγισε ως το πιο σχετικό με την επερώτηση q.

Α. Υπολογίστε την Ακρίβεια, την Ανάκληση και το F-measure για τα 2 συστήματα.

Β. Να σχεδιάσετε τις καμπύλες Ακρίβειας-Ανάκλησης για τα 2 συστήματα.

**Α.**

**(S1): R(S1) ΣΑ/Α**=8/10=80% **P(S1)** 8/15=53%

**F-measure : 2 PR / (P+R):** 0,848/1,33=63,7%

**(S2): R(S2)** **ΣΑ/Σ** = 9/10=90% **P(S2)** 9/15=60%

**F-measure:** **2 PR / (P+R)** =1,08/1,5=72%

**Β.**

**(S1,q)**

P=1/1=100% R=1/10=10%

P=2/3=66% R=2/10=20%

P=3/4=75% R=3/10=30%

P=4/9=44% R=4/10=40%

P=5/11=45% R=5/10=50%

P=6/12=50% R=6/10=60%

P=7/13=54% R=7/10=70%

P=8/14=57% R=8/10=80%

**(S2,q)**

P=1/1=100% R=1/10=10%

P=2/2=100% R=2/10=20%

P=3/4=75% R=3/10=30%

P=4/6=66% R=4/10=40%

P=5/8=62% R=5/10=50%

P=6/9=66% R=6/10=60%

P=7/13=53% R=7/10=70%

P=8/14=57% R=8/10=80%

P=9/15=60% R=9/10=90%