PROBLEM #1

(a) Cumulative Gain:- (CG)

for q_1:

$$CG = \sum_{i=1}^{7} rel_i = 2+0+2+2+1+1+1 = 9$$

for q_2:

$$CG = \sum_{i=1}^{7} rel_i = 2+2+2+1+1+0+0 = 8$$

(b) Discount Cumulative Gain:-

for q_1

i	doc	rel_i	$\log_2(i+1)$	$rel_i/\log_2(i+1)$
1	4	2	1	2
2	5	0	1.585	0
3	6	2	2	1
4	7	2	2.322	0.861
5	1	1	2.585	0.3868
6	8	1	2.807	0.3562
7	9	1	3	0.333

$$DCG = \sum_{i=1}^{7} \frac{rel_i}{\log_2(i+1)} = 2+0+1+0.861+0.3868+0.3562+0.333 = 4.9372 \approx 5$$

Ans

for q2: DCG

i	doc	reli	$\log_2(i+1)$	$rel_i/\log_2(i+1)$
1	2	2	1	2
2	3	2	1.525	1.261
3	8	2	2	1
4	4	1	2.322	0.430
5	5	1	2.585	0.3868
6	9	0	2.807	0
7	1	0	3	0

$$DCG = \sum_{i=1}^{7} \frac{rel_i}{\log_2(i+1)} = 2 + 1.261 + 1 + 0.430 + 0.3868 + 0 + 0 = 5.0778$$

(c) Normalized DCG:

for q1:

$$IDCG = \sum_{i=1}^{7} \frac{rel_i}{\log_2(i+1)} = \frac{2}{1} + \frac{2}{1.585} + \frac{2}{2} + \frac{1}{2.322} + \frac{0}{2.585} + \frac{0}{2.807} + \frac{0}{3}$$

$$= 5.435$$

$$nDCG = \frac{DCG}{IDCG} = \frac{4.937}{5.435} = 0.908$$

≤ 935

for q2:

IDCG = ~~$\frac{rel_1}{\log_2(1)}$~~ $+ \frac{rel_2}{\log_2(2)}$

$= \frac{2}{1} + \frac{2}{1.585} + \frac{2}{2} + \frac{1}{2.322} + \frac{1}{2.585} + 0 + 0$

$= 5.0778$

$nDCG_1 = \frac{DCG_1}{IDCG} = \frac{1}{IDCG}$

PROBLEM # 2

		Judge 2		
		Yes = 1	No = 0	Total
Judge One	Yes = 1	2	2	4
	No = 0	2	2	4
	Total	4	4	8

(a) Kappa Measure:—

$$K = \frac{P(A) - P(E)}{1 - P(E)}$$

Observed proportions of the times the judges agreed:

$$P(A) = \frac{2+2}{8} = \frac{1}{2} = 0.5$$

P(non rel) = 4+4 / 8+8 = 0.5
P(Rel) = 4+4 / 8+8 = 0.5

$P(E) = P(\text{rel})^2 + P(\text{non Rel})^2$
$= (0.5)^2 + (0.5)^2$
$= 0.5$

$$K = \frac{0.5 - 0.5}{1 - 0.5} = 0$$

If two judges agree:-

(b) Precision: $\frac{2}{5} = 0.4$

Recall: $\frac{2}{2} = 1$

$F_1 = \frac{2PR}{P+R} = \frac{2(0.4)(1)}{0.4+1} = 0.5714$

$\beta = 1$

(c) If either judge agrees

Precision $\frac{4}{5} = 0.8$

Recall $= \frac{4}{6} = 0.67$

$F_1 = \frac{2PR}{P+R} = \frac{2(0.8)(0.67)}{(0.8)(0.67)} = 0.729$

$\beta = 1$

A

PROBLEM # 3

Roechio Algorithm:-

$$\vec{q_m} = \alpha \vec{q_0} + \frac{\beta 1}{|D_r|} \sum_{\vec{d_j} \in D_r} \vec{d_j} - \frac{\gamma 1}{|D_{nr}|} \sum_{\vec{d_j} \in D_{nr}} \vec{d_j}$$

Sum of relevant Doc Vectors

$$\sum_{\vec{d_j} \in D_r} \vec{d_j} = \langle 0.61, 0.71, 0.31, 0.21, 0.55, 0.81, 0.4 \rangle$$

Sum of non relevent vectors

$$\sum_{\vec{d_j} \in D_{nr}} \vec{d_j} = \langle 0.11, 0.21, 0.11, 0.11, 0.12, 0.12, 0.01 \rangle$$

$$\alpha = 0.1, \beta = 0.2, \gamma = 0.4$$

$$\vec{q_m} = (0.1)(\langle 0.01, 0.22, 0.11, 0.01, 0.01, 0.22, 0.1 \rangle) + \frac{(0.2)}{3}(\langle 0.61, 0.71, 0.31, 0.21, 0.55, 0.81, 0.41 \rangle) - \frac{(0.4)}{1}(\langle 0.11, 0.21, 0.11, 0.11, 0.12, 0.12, 0.01 \rangle)$$

$$\vec{q_m} = \langle -3 \times 10^{-3}, -0.015, -0.013, -0.029, -0.0103, 0.02, 0.03347 \rangle$$

(b) $\alpha = 0.1$, $\beta = 0.2$, $\gamma = 6$

$$\vec{q_m} = (0.1)\langle 0.01, 0.22, 0.11, 0.01, 0.01, 0.22, 0.17\rangle + \frac{(0.2)}{3}\langle 0.61, 0.71, 0.31, 0.21, 0.55, 0.81, 0.417\rangle$$

$$= \langle 0.041, 0.069, 0.031, 0.015, 0.0377, 0.072, 0.03747\rangle$$

(c) When

i) $\alpha = 1$ and $\frac{\beta}{|D_r|}\sum_{d_j \in D_r} \vec{d_j} = \frac{\delta}{|D_{nr}|}\sum_{d_j \in D_{nr}} \vec{d_j}$

q_m will be same as original q.

ii) No, if β is very small and γ is very large q_0 may be closer to centroid of relevant documents