National University of Computer & Emerging Sciences
FAST-Karachi Campus
CS4051- Information Retrieval
Quiz#2

Dated: March 14, 2023   Marks: 20

Time: 20 min.

Std-ID: _____Sol_____

Question No. 1

What are the issues arising when we use evaluation matrices developed for unranked retrieval to evaluate ranked retrieval? **[5]**

Unranked Retrieval systems return all relevant document at the same level (Flat Results) while the ranked retrieval systems return a rank ordering of the results that top documents are more relevant hence the user favor this type of retrieval. The evaluation of unranked retrieval like Precision and Recall are computed using set-oriented measure and hence does not suitable for ranked results. These measure simply fails to incorporate the ranking order of the results – returned by rank retrieval systems.

Question No. 2

What is a Test Collection in the context of Information Retrieval? What are its essential components? **[5]**

To measure ad hoc information retrieval effectiveness in the standard way, we need a test collection consisting of three things:

1. A document collection

2. A test suite of information needs, expressible as queries

3. A set of relevance judgments, standardly a binary assessment of either relevant or non-relevant for each query-document pair.

Question No. 3

Assume there are 16 total relevant documents in a collection. Consider the following result against a query that return 10 documents in the given order., where plus indicates relevant and minus indicates non-relevant:

{+, +, −, −, −, −, −, +, +, +}

Calculate the following evaluation measures on the ranked list.? **[2.5 x 4]**

a. **What is the Precision?**

   Precision = # relevant retrieved / total retrieved = 5/10 = 0.5

b. **What is the Recall?**

   Recall = # relevant retrieved / total relevant = 5/16 = 0.312

c. **What is the F1 Score?**

   F1= (2 *P*R)/ (P+R) = (2*0.5*0.312)/(0.5+0.312)= 0.312/0.812= 0.384

d. **What is Average Precision?**

   AP = (1+1+ 3/8+4/9+5/10) / 5 = 3.319 / 5 = 0.66