


Name _____
Section _____

Roll No _____

National University of Computer and Emerging Sciences, Lahore Campus

	Course:	Information Retrieval	Course Code:	CS317
	Program:	BS(Computer Science)	Semester:	Fall 2018
	Duration:	25 Minutes	Total Marks:	14
	Paper Date:	7-Nov-18	Weight	3.3%
	Section:	B	Page(s):	2
	Exam:	Quiz 2	Roll No:	

Question1:

What are advantages of using Average precision over Precision @k ? Explain with examples. [2 Marks]

Solution:

1. $P@k$ = user centric, AP is system centric
2. AP evaluates quality of entire rank list
3. AP is sensitive to every change in rank list. List 1 has better quality s compared to list 2. AP is higher for list 1, while $P@5$ is same for both lists

List 1

R
R
N
N
N

List 2

N
N
N
R
R

Question2:

Consider following collection of 3 documents.

Document	Words
D1	a b b a b b c
D2	a a b a b a
D3	b b b b b b c c

Query = < a b >

Use Witten Bell smoothing to find similarity of document D1 with query.

Solution:

$N = 7, V = 3$

$\text{Prob}(a) = 7/10(2/7) + 3/10(6/21) = 0.29$

$\text{Prob}(b) = 7/10(4/7) + 3/10(12/21) = 0.57$

$\text{Score f D1} = 0.29 * 0.57 = 0.17$

Question3:

How does Rocchio Feedback improve retrieval performance?

Solution:

Query vector is modified so that it becomes closer to relevant documents and moves away from non-relevant documents. The centroid vector of relevant documents is added in vector of original query and centroid vector of non-relevant documents is subtracted from vector of original query. The modified query vector has higher similarity with relevant documents and less similarity with non-relevant documents.