

National University of Computer & Emerging Sciences  
FAST-Karachi Campus  
Information Retrieval (CS317)  
Quiz#2

Marks: 20

Std-ID: \_\_\_\_\_ Sol \_\_\_\_\_

### What are some of the limitations of Vector Space Model (VSM) for IR? [5]

- ### Question No.2

**What are the two assumptions on which Rocchio's algorithms is effective? [5]**

1. The user has sufficient knowledge about the initial query and associated documents.
2. The relevance prototypes are well balance and well behaved. This means terms and their distribution in relevant documents are same and these are different from the terms and their distribution for non-relevant documents. Both relevant and non-relevant documents are well separated.

### Question No.3

Comments on the following statement as TRUE or FALSE with justification (1-2 line explanations). 5 X 2 marks each.

**1. Average precision cannot discriminate a rank order search results.**

TRUE. Average precision is average of the precision level values at different order of the results. Hence a precision list 1,2,3,4, and 4,3,2,1 may produce average precision of 2.5 in each case. Therefore, it cannot discriminate the ordering of precision in the list.

**2. Normalized DCG metric does not penalize for bad/missing documents in the result.**

TRUE. NDCG only works for the relevant documents in the ranked results. It does not consider the bad document or missing document from the retrieval.

**3. Information retrieval performance can be improved with multiple rounds of relevance feedback.**

FALSE. Only the first round of relevance feedback shows identifiable improvement.

**4. Precision is affected by making false positive error on retrieval.**

TRUE. Precision is computed as **relevant\_retrieved** over **total\_retrieved**. If we retrieved a false positive (non-relevant document) the denominator increases and thus decreases the precision.

**5. Relevance feedback is of no value to web-search.**

TRUE. Web-search is a precision critical application. Relevance Feedback(RF) generally improve recall, improving recall may not show any appreciation. On the other hand, web search need quick response and RF may delay the result.