Experiment No. 3 (Group - B)

Page No.
Date

Problem Statement:
Implement a program to calculate precision & recall for sample input.

Objectives:
1. To understand precision & recall in information retrieval.
2. To study Indexing structures for informatn retrieval.

Theory:
Precision & Recall in Information Retrieval.

Informatn Retrival system can be measured with two metrices: precision & recall. It can be divided into 4 categories.
1) Relavant & Retrieved
2) Relavant & not retrieved
3) non-relavant & Retrieved
4) Non-relevant & non retrieved.

Relevant items are those documents that help the user in answering his question. Non relavent items are items that don't provide actually useful information. for each item there are two possibilities: it can be retrieved by the user's query. Precision is defined as the ration of the no. of relevant & retrieved documents.
- Recall is defined as ratio of the no. of retrieved & relavent documents to the possible relevant documents.

Page No.
Date

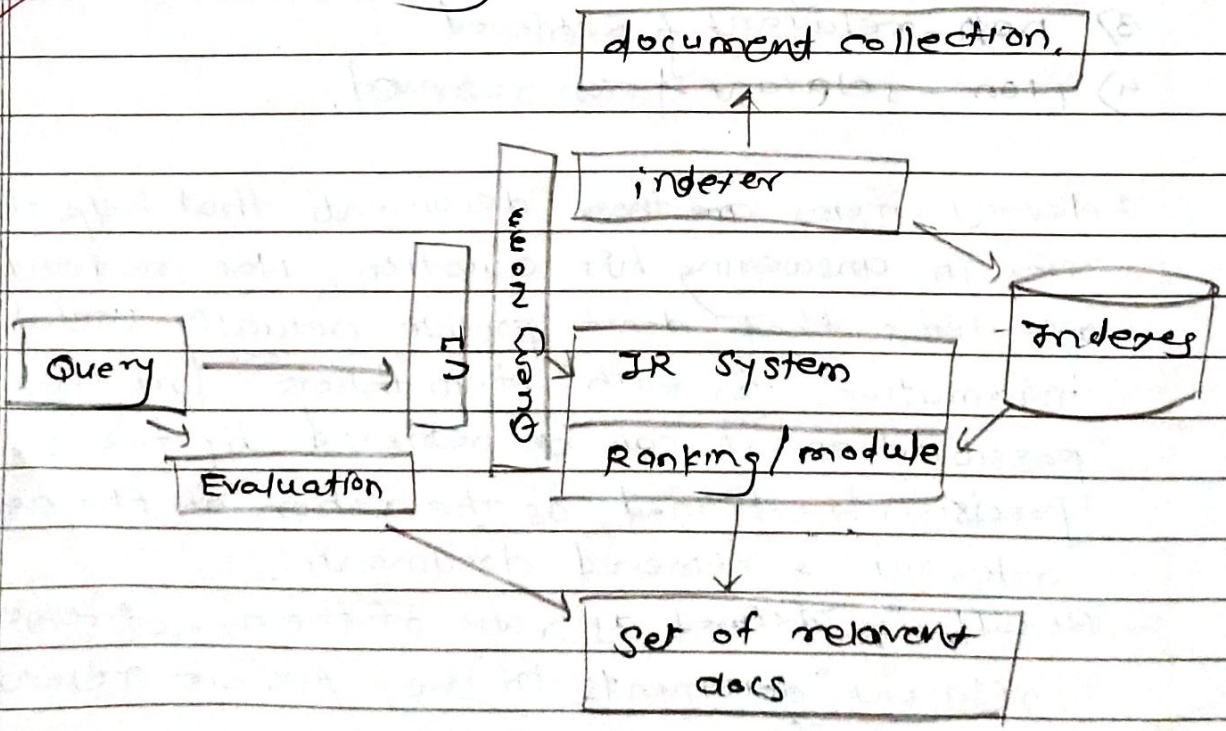
Precision :- $\frac{\#(\text{relevant item retrieved})}{\#(\text{retrieved items})}$

Recall = $\frac{\#(\text{relevant item retrieved})}{\#(\text{relavent items})}$

		The truth relavent	Non relevant
What the system thinks	Retrived	TP	FP
	Non retrieved	FN	TN

Precision / Recall trade-off:

You can increase recall by returning more docs. Recall is a non-decreasing fun^n of the no. of docs retrieved.



Page No.
Date

Precision & Recall explanation:

Consider,

I: an information request

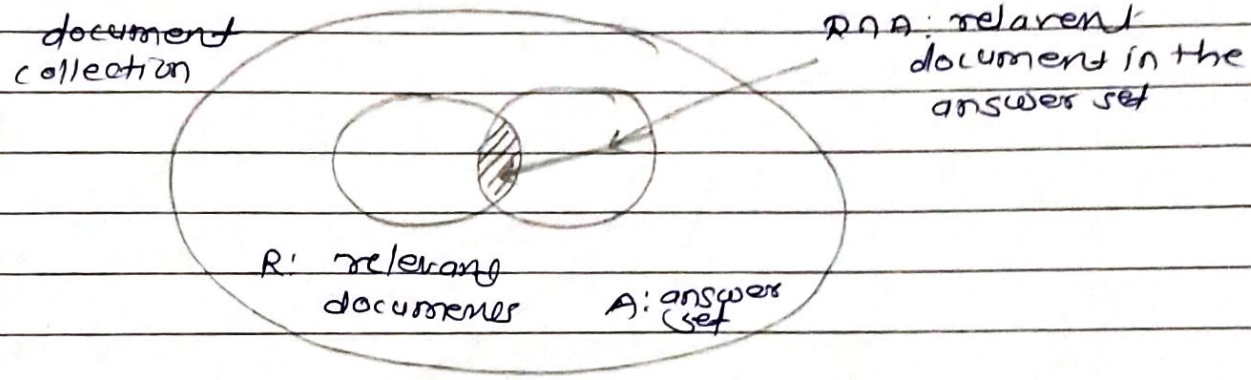
R: the set of relavent documents for I

A: the answer for I, generated by IR system

$R \cap A$: the intersection of the sets R & A

$|A|$ - no. of docs in the set A

$|Ra|$ - no. of docs in the intersection of sets R & A



Recall: $\frac{|R \cap A|}{|R|}$

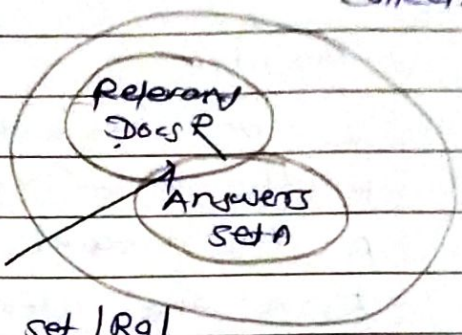
Precision = $\frac{|R \cap A|}{|A|}$

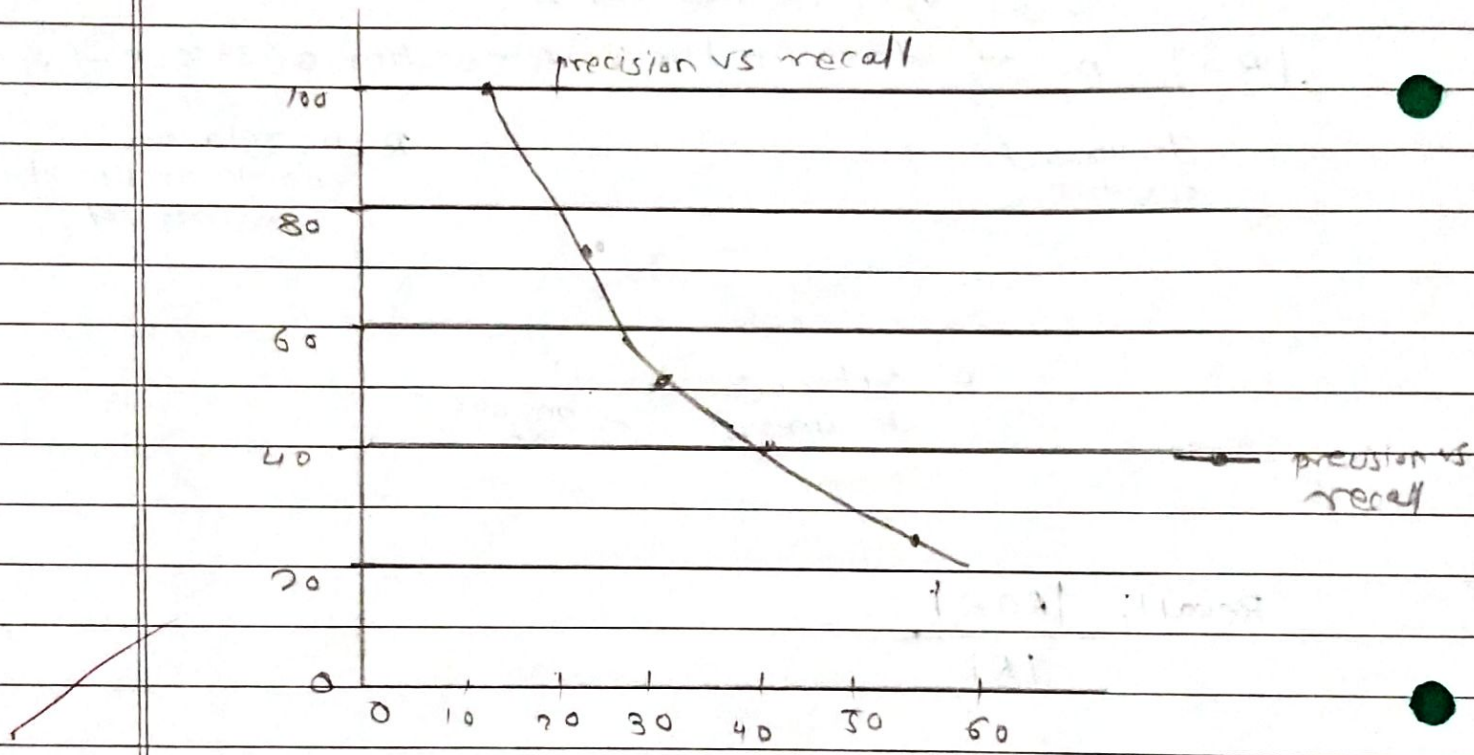
Recall: R = $\frac{\text{no. of relavent items retrieved}}{\text{Total no. of relevant items in collection}}$

Precision: P = $\frac{\text{no. of relevant items retrieved}}{\text{Total no. of items retrieved}}$

$$\text{Recall} = \frac{|Rq|}{|R|}$$

$$\text{precision} = \frac{|Rq|}{|A|}$$





Conclusion :- Implementation is concluded by executing a program to calculate precision and recall for sample input with relavent documents Rq1 & for query q2

28
21-10-2024