1. Calculation of Evaluation Measures

1.1 Calculate Precision, Recall and 𝐹1 -Measure for all three result sets, treating them as unranked result sets.

***Precision*** = =

***Recall*** *=*  =

**-*Measure*** *=*

First result set:

*Precision* =  0.43;

*Recall* = = 0.75;

-*Measure =*  0.55

Second result set:

*Precision* =  *=* 0.5;

*Recall* =  *=* 0.5;

-*Measure =* = 0.5

Third result set:

*Precision* =  *=* 0.25;

*Recall* = 1;

-*Measure =* = 0.4

1.2 Calculate Precision at 𝑘 (P@𝑘) with 𝑘 = 1 and 𝑘 = 5, and R-Precision for all three result sets, treating them as ranked result sets.

**Precision at 𝑘** correspondents to the number of relevant results on the first search results page[[1]](#footnote-1)

**R-precision** requires knowing all documents that are relevant to a query. The number of relevant documents R, is used as the cutoff for calculation, and this varies from query to query. Precision is equal to recall at the R-th position.[[2]](#footnote-2)

First result set:

|  |  |  |
| --- | --- | --- |
| k | #relevant docs | P@k |
| 1 | 0 | 0 |
| 5 | 3 | 0.600 |

*R-Precision* =

Second result set:

|  |  |  |
| --- | --- | --- |
| k | #relevant docs | P@k |
| 1 | 1 | 1 |
| 5 | 2 | 0.400 |

*R-Precision* =

Third result set:

|  |  |  |
| --- | --- | --- |
| k | #relevant docs | P@k |
| 1 | 0 | 0 |
| 5 | 1 | 0.200 |

*R-Precision* =

1. F-Measure

2.1 Give the definition of the more-general 𝐹𝛽 -score for arbitrary values of 𝛽.

F-measure - a single measure that trades off precision versus recall, which is the weighted harmonic mean of precision and recall[[3]](#footnote-3):

, where values of β < 1 emphasize precision, while values of β > 1 emphasize

2.2 Explain why the choice of 𝛽 = 1 is most common.

Because the default balanced F-measure equally weights precision and recall, which means making 𝛽 = 1.

2.3 Explain how the behavior of the measure would change for 𝛽 = .

With 𝛽 = the F-measure will emphasize precision more than recall, in other words we’re weighting times as much importance to recall as precision[[4]](#footnote-4).

1. https://en.wikipedia.org/wiki/Evaluation\_measures\_(information\_retrieval)#Precision\_at\_K [↑](#footnote-ref-1)
2. https://en.wikipedia.org/wiki/Evaluation\_measures\_(information\_retrieval)#R-Precision [↑](#footnote-ref-2)
3. https://nlp.stanford.edu/IR-book/pdf/08eval.pdf [↑](#footnote-ref-3)
4. http://www.dcs.gla.ac.uk/Keith/Chapter.7/Ch.7.html [↑](#footnote-ref-4)