Français (https://moodle.epfl.ch/mod/quiz/review.php?attempt=15298&lang=fr)
English (https://moodle.epfl.ch/mod/quiz/review.php?attempt=15298&lang=en)

Started on	Thursday, 8 March 2018, 10:45
State	Finished
Completed on	Thursday, 8 March 2018, 11:00
Time taken	14 mins 54 secs
Grade	**5.00** out of 8.00 (**63**%)

Question 1

Correct

Mark 1.00 out of 1.00

Flag question

Which of the following is **true**?

Select one:

- a. High recall hurts precision ✓
- b. High recall implies low precision
- c. High precision implies low recall
- d. High precision hurts recall

The correct answer is: High precision hurts recall

Question 2

Correct

Mark 1.00 out of 1.00

Flag question

Which of the following is an advantage of Vector Space Retrieval model?

Select one:

- a. Allows to retrieve documents that do not contain any of the query terms
- b. No theoretical justification is needed why the model works
- c. Produces provably correct query results
- d. Enables ranking of query results according to cosine similarity function ✓

The correct answer is: Enables ranking of query results according to cosine similarity function

Question 3

Incorrect

Mark 0.00 out of 1.00

Flag question

Recall can be defined as:

Select one:

- a. P(retrieved documents | number of documents)
- b. P(relevant documents | number of documents) ✗

c. P(retrieved documents | relevant documents)

d. P(relevant documents | retrieved documents)

The correct answer is: P(retrieved documents | relevant documents)

Question 4

Correct

Mark 1.00 out of 1.00

Flag question

Which one of the following is **wrong**. Schema mapping is used to:

Select one:

a. Support schema evolution of databases

b. Optimize the processing of queries ✓

c. Reconcile different logical representations of the same domain

d. Overcome semantic heterogeneity

The correct answer is: Optimize the processing of queries

Question 5

Correct

Mark 1.00 out of 1.00

Flag question

Data being classified as unstructured or structured depends on the:

Select one:

a. Type of physical storage

b. Amount of data

c. Degree of abstraction ✓

d. Level of human involvement

The correct answer is: Degree of abstraction

Question 6

Incorrect

Mark 0.00 out of 1.00

Flag question

In a Ranked Retrieval result, the result at position k is non-relevant and at k+1 is relevant. Which of the following is always true (P@k and R@k are the precision and recall of the result set consisting of the k top ranked documents)?

Select one:

a. P@k-1 = P@k+1

b. R@k-1 = R@k+1

c. P@k-1 > P@k+1 ✗ Try with k-1 at 1/2

d. R@k-1 < R@k+1

The correct answer is: R@k-1 < R@k+1

Question 7

Incorrect

Mark 0.00 out of 1.00

Flag question

The number of non-zero entries in a column of a term-document matrix indicates:

Select one:

a. how relevant a term is for a document

b. how many terms of the vocabulary a document contains ✗

c. none of the other responses is correct because of tf-idf

d. how often a term of the vocabulary occurs in a document

The correct answer is: none of the other responses is correct

Question 8

Correct

Mark 1.00 out of 1.00

Flag question

Thang, Jeremie and Tugrulcan have built their own search engines. For a query Q, they got precision scores of 0.6, 0.7, 0.8 respectively. Their F1 scores (calculated by same parameters) are same. Whose search engine has a higher recall on Q?

Select one:

a. Tugrulcan

b. Jeremie

c. We need more information

d. Thang ✓ if F1 score is the same, then higher recall is needed to "balance" low precision

The correct answer is: Thang

Finish review (https://moodle.epfl.ch/mod/quiz/view.php?id=977177)