

Started on	Thursday, 8 March 2018, 10:44 AM
State	Finished
Completed on	Thursday, 8 March 2018, 10:59 AM
Time taken	15 mins 6 secs
Grade	7.00 out of 8.00 (88%)

Question 1

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 ✓

The correct answer is: Thang

Question 2

Correct

Mark 1.00 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. $R@k-1 = R@k+1$
- ☒ b. $R@k-1 < R@k+1$ ✓
- ☐ c. $P@k-1 = P@k+1$
- ☐ d. $P@k-1 > P@k+1$

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

Question 3

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

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

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. Overcome semantic heterogeneity
- ☐ b. Support schema evolution of databases
- ☐ c. Reconcile different logical representations of the same domain
- ☒ d. Optimize the processing of queries ✓

The correct answer is: Optimize the processing of queries

Question 5

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 precision hurts recall
- ☐ c. High precision implies low recall
- ☐ d. High recall implies low precision

The correct answer is: High recall hurts precision

Question 6

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. none of the other responses is correct
- ☐ c. how often a term of the vocabulary occurs in a document
- ☒ d. how many terms of the vocabulary a document contains ❌

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

Question 7

Correct

Mark 1.00 out of 1.00

🚩 Flag question

Data being classified as unstructured or structured depends on the:

Select one:

- ☒ a. Degree of abstraction ✔️
- ☐ b. Level of human involvement
- ☐ c. Type of physical storage
- ☐ d. Amount of data

The correct answer is: Degree of abstraction

Question 8

Correct

Mark 1.00 out of 1.00

🚩 Flag question

Recall can be defined as:

Select one:

- ☐ a. $P(\text{relevant documents} \mid \text{number of documents})$
- ☐ b. $P(\text{relevant documents} \mid \text{retrieved documents})$
- ☐ c. $P(\text{retrieved documents} \mid \text{number of documents})$
- ☒ d. $P(\text{retrieved documents} \mid \text{relevant documents})$ ✔️

The correct answer is: $P(\text{retrieved documents} \mid \text{relevant documents})$