Among the following groups of concepts, which group is not totally relevant to a search engine?

○ A. word stemming, hashing, compression

◉ B. inverted file index, stop words, precision

○ C. distributed index, backtracking, query

○ D. posting list, thresholding, recall

C

Which of the following is NOT a step in the process of building an inverted file index?

◉ A. Get the posting list for each term and calculate the precision

○ B. Read in strings and parse to get words

○ C. Check dictionary with each term: if it is not in, insert it into the dictionary

○ D. Use stemming and stop words filter to obtain terms

答案正确：1分　♀ 创建提问 ☑

While accessing a term stored in a B+ tree in an inverted file index, range searchings are expensive.

F

In general, comparing with the terms with short posting lists, those with very long posting lists are less important.

◉ T　○ F

答案正确：1分　♀ 创建提问 ☑

Inverted file contains a list of pointers to all occurrences of a term in the text.

◉ T　○ F

答案正确：1分　♀ 创建提问 ☑

In a search engine, thresholding for query retrieves the top k documents according to their weights.

F

When measuring the relevancy of the answer set, if the precision is high but the recall is low, it means that most of the relevant documents are retrieved, but too many irrelevant documents are returned as well.

F

For the document-partitioned strategy in distributed indexing, each node contains a subset of all documents that have a specific range of index.

○ T ○ F

答案正确: 1 分   ♡ 创建提问 ☒

Precision measures the quality of all the retrieved documents.

T

Word stemming is to eliminate the commonly used words from the original documents.

F

For the document-partitioned strategy in distributed indexing, each node contains a subset of all documents that have a specific range of index.

○ T ○ F

答案正确: 1 分   ♡ 创建提问 ☒