

TD Map Reduce

Ludovic Denoyer

19 avril 2013

1 Rappel de Cours / Map Reduce

Question 1

Rappelez les principes de map-reduce ; quel intérêt à une telle méthode.

Question 2

Donnez la définition de **map**

Question 3

Donnez la définition de **reduce**

Question 4

Rappelez les principes de l'implémentation de map-reduce sous MongoDB

2 Rappel de Cours / TF-IDF

Question 1

Rappelez les principes d'un moteur de recherche

Question 2

Comment se calcule le score $RSV(d, q)$?

Question 3

Donnez la définition de **TF**

Question 4

Donnez la définition de **IDF**

3 Exercice Map-Reduce et TF-IDF

Question 1

Définir une fonction map-reduce permettant le calcul du DF de chaque mot

Question 2

Définir une fonction map-reduce permettant le calcul du TF de chaque mot/document (la position du mot sera ignorée, on réverra la fréquence des mots dans le document)

Question 3

Définir le schéma SQL de stockage. Quel est la requête permettant de calculer les scores des documents pour une requête donnée ?

Question 4

Implémentez en JAVA la création de l'index des IDF, et de l'index inversé

Question 5

Implémentez la fonction de score en JAVA, et le service de recherche associé