
TD 3 : MapReduce

Exercice 1 :

Soit un grand document de type texte. Nous souhaitons réaliser quelques analyses sur son contenu. Plus particulièrement, catégoriser les mots en 4 catégories :

- Minuscule : 1 lettre ;
- Petit : 2 à 4 lettres ;
- Moyen : 5 à 9 lettres ;
- Grand : plus de 10 lettres

Travail à faire :

- Donner le schéma détaillé des opérations Map-Reduce
- Ecrire en pseudo code les fonctions Map et Reduce permettant de retourner le nombre de mots dans chaque catégorie.
- Traduire le pseudos codes en langage java.

Exercice 2 :

On reçoit d'un fournisseur une livraison de fruits (pommes, ananas, oranges, etc.). On décide d'effectuer un test qualité en leur donnant 0 ou 1 pour les critères suivants (trop mûr, taché, déformé, etc). Pour chaque espèce de fruit, on veut calculer la proportion de fruits trop mûrs. Comment faire avec MapReduce?

- Présentez avec explication) les taches Map-Reduce permettant de retourner le résultat souhaité à l'aide d'un schéma détaillé.
- Ecrivez en pseudo-code (algorithmes) les fonctions Map et Reduce.
- On veut faire le même calcul pour tous les critères. Comment faire ?