

CES Data Scientist - PageRank Exam

Youcef KACER (youcef.kacer@telecom-paristech.fr)

26 septembre 2016

Exercise 1 : We expose here after the system of linear equations that solve PageRank scores associated to graph G (r_i being PageRank score for page i) :

$$\begin{aligned}r_1 &= r_4 \\r_2 &= 1/2 * r_1 \\r_3 &= 1/2 * r_1 + r_2 \\r_4 &= r_3 \\1 &= r_1 + r_2 + r_3 + r_4\end{aligned}$$

We can solve manually the system :

$$\begin{aligned}r_1 &= r_3 \\r_1 &= r_4 \\r_2 &= 1/2 * r_1 \\1 &= 7/2 * r_1\end{aligned}$$

$$\begin{aligned}r_1 &= 2/7 \\r_2 &= 1/7 \\r_3 &= 2/7 \\r_4 &= 2/7\end{aligned}$$

Otherwise, we can solve the system using previous PageRank lab using MapReduce. We just need to adapt *edge_list.txt* to the graph G . In this file, each line represents a page indice and the page indices it links to :

```
1 2 3
2 3
3 4
4 1
```

Exercise 2 : bestiole qui fait bloub.

Exercise 3 : bestiole qui fait rire.

Exercise 4 : bestiole qui fait mal.