

# PageRank

## Planteamiento del problema

En esta práctica se espera obtener el PageRank de los nodos del grafo obtenido de la página de facebook de Vox.

## Procedimiento

Utilizando los datos del grafo de Vox que se utilizó en la práctica anterior, así como el código utilizado, en esta ocasión se procederá a implementar en la clase Graph el algoritmo de PageRank, y en la clase Node, alojar el PageRank que se va obteniendo.

## Desarrollo

En la clase Graph, se eliminaron los dos algoritmos utilizados en la práctica anterior (BFS y DFS) para agregar el la función del PageRank en la clase Graph y Otra para alojar el valor del PageRank en cada nodo, es decir, actualizar su valor.

Clase Graph:

```
def page_rank(self, iteraciones = 0):
    keys = self.nodes.keys()
    for _ in range(iteraciones):
        for key in keys:
            self.nodes[key].page_rank_value(d=0.85)
    page_rank_d={}
    for key in keys:
        page_rank_d[self.nodes[key].name] = self.nodes[key].page_rank
    key_value = zip(page_rank_d.keys(), page_rank_d.values())
    return dict(sorted(key_value, key=lambda x: x[1], reverse=True))
```

Clase Node:

```
def page_rank_value(self, d):
    acumulado = 0
    for node in self.adjacencyInverse.values():
        acumulado += node.page_rank / node.connected
    self.page_rank = (1-d) + (d) * acumulado
```

## Resultados

Una vez ejecutado el algoritmo pude observar que el nodo con el mayor PageRank es la **NBA**. Sin embargo, al momento de analizar cuál es el nodo menos influyente, me encontré con que existen varios nodos con el mismo PageRank que teoría son el menor.

Page Rank	
NBA	6.638976624662659
SB Nation	5.213629864141949
Boston Celtics	3.0008600580520195
Sprint	2.993395535356116
Baserange	2.2546030715320455
Yahoo Sports NBA	2.2438433947231973
Totokaelo	2.13481402083005
Puck Daddy	2.0335447376582176
UFC	1.763547811578075
Dr. Saturday	1.3837571073815451
Uproxx Hip-Hop	1.3078206626623832
UPROXX News	1.2834779887940693
SB Nation Combat Sports	1.073882016958538
Bleeding Green Nation: For Philadelphia Eagles Fans	1.0132064500933609
Racked	0.997273541177781
theScore NBA	0.9665886834034486
The Verge	0.9568575498197314
Vox	0.9140117794458956
Curbed	0.8818569875135075
Yahoo Sports	0.8456195717702744
Vox Media, Inc	0.8297815719945614
Sports Illustrated	0.8196479710705195
BloodyElbow.com	0.7443572538476655
Theory	0.7122646137541826

  

SB Nation Seattle	0.17183046987448602
The Nova Blog	0.17183046987448602
SB Nation St. Louis	0.17183046987448602
SB Nation Houston	0.17183046987448602
SB Nation Bay Area	0.17183046987448602
La Perla Lingerie	0.17018291690478846
Journelle	0.17018291690478846
BARE NECESSITIES	0.17018291690478846
Adore Me	0.17018291690478846
Agent Provocateur	0.17018291690478846
UNIQLO USA	0.17018291690478846
ZARA	0.17018291690478846
Topshop	0.17018291690478846
The RealReal	0.17018291690478846
THE OUTNET.COM	0.17018291690478846
SSENSE	0.17018291690478846
Pixie Market	0.17018291690478846
Madewell	0.17018291690478846
Aritzia	0.17018291690478846
H&M	0.17018291690478846
Farfetch	0.17018291690478846
& Other Stories	0.17018291690478846
COS	0.17018291690478846
Barneys Warehouse	0.17018291690478846
ASOS	0.17018291690478846
Anthropologie	0.17018291690478846
Reformation	0.17018291690478846
Bath & Body Works	0.17018291690478846
Le Labo	0.17018291690478846
Gap	0.17018291690478846

→ practica2-2 git:(master) x █

## Conclusiones

Al finalizar esta práctica hay algo que me llamó mucho la atención. Para empezar, en comparación con la práctica anterior, el nodo con la mayor cantidad de hijos, no implica que sea también el más influyente al aplicar este algoritmo. También, recalcar que ahora

nos mostró otros resultados en cuanto al nivel de influencia en comparación con los que realizamos al principio del semestre.

Otra cosa curiosa es que existen muchos nodos que tienen el mismo PageRank, en especial los últimos nodos.