



Análise de Links

Agente Educacional
Marcelo Pita

Agenda

Tarefa de análise de links

PageRank

HITS



Tarefa de **Análise de Links**

Visa **extrair conhecimento de alta ordem** da rede baseado na importância relativa dos nós, seu **isolamento** ou **similaridade**.

Tipos de tarefas

- *Link-based Object Classification (LOC)*

Atribuir rótulos aos nós de acordo com suas conectividades

- *Link-based Object Ranking (LOR)*

Avaliar nós de acordo com sua importância relativa na rede

- *Link prediction*

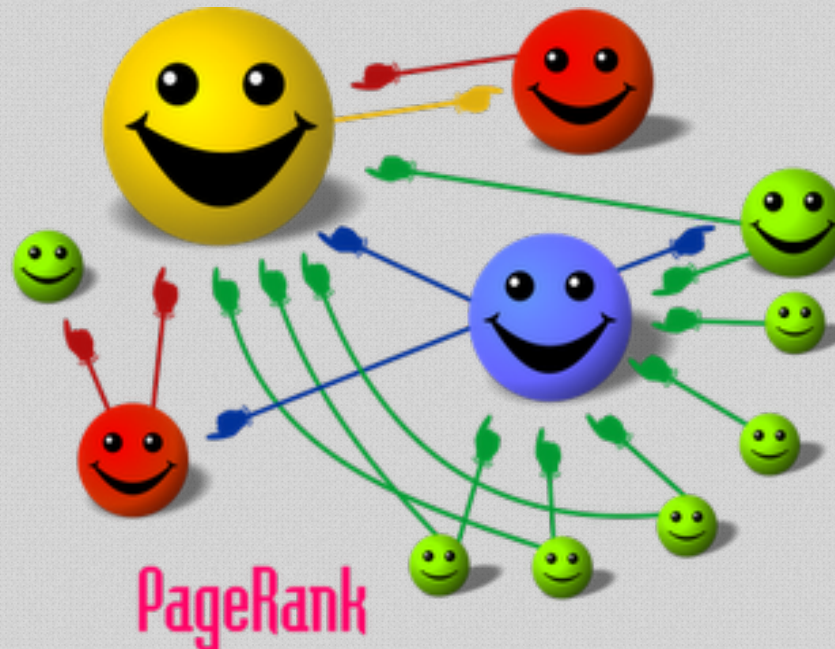
Deduzir novos links plausíveis na rede

Vamos nos concentrar aqui



PageRank

No algoritmo *PageRank* um nó é mais importante se é apontado por outros nós importantes na rede.



Algoritmo *PageRank*

Em uma rede com n nós, atribuímos o mesmo *PageRank* inicial de valor $1/n$ para todos os nós.

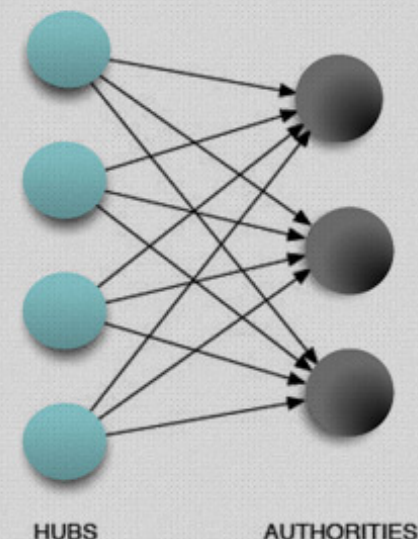
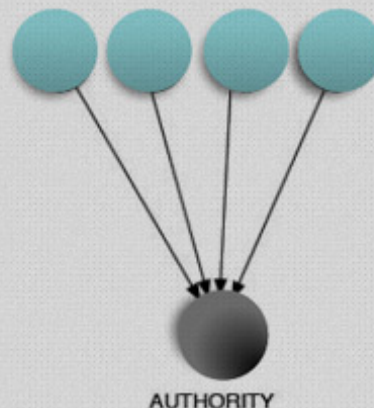
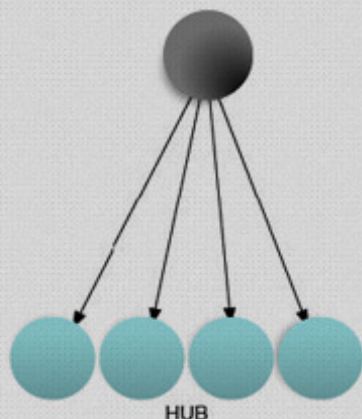
Executar k passos

- Cada nó divide igualmente seu *PageRank* atual para todos os nós que ele aponta. Se o nó não aponta para nenhum outro, então ele passa o *PageRank* para si
- Cada nó atualiza seu *PageRank* como a soma dos *PageRanks* que ele recebe dos nós que o apontam

HITS (*Hyperlink Induced Topic Search*)

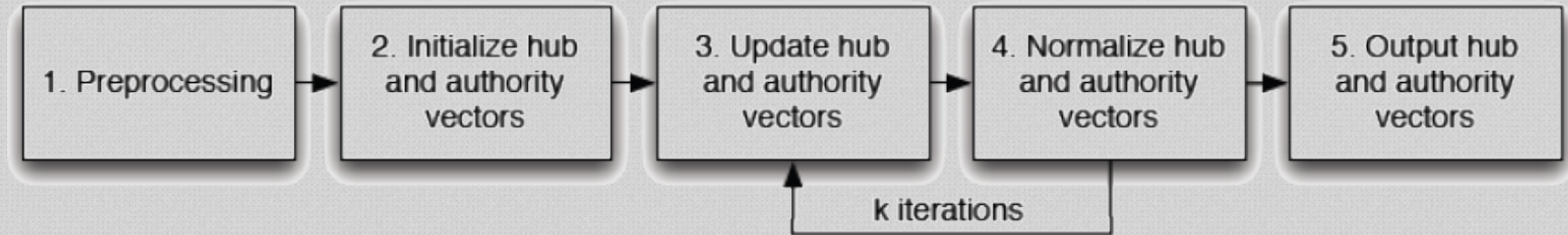
Para **grafos direcionados**, originalmente projetado para verificar a importância de páginas Web sob **duas perspectivas**

- Qualificação como **hub**
- Qualificação como **autoridade**



Algoritmo HITS

Algoritmo iterativo para calcular *hub* e *autoridade* de cada nó da rede



- Inputs

- Grafo direcionado (matriz de adjacência)
- Número de iterações

- Outputs

- Vetores de pontuação de *hub* e autoridade

- 1 - Todos os scores de *hub* e *autoridade* são iniciados com 1
- 2 - Executar *k* atualizações de scores e *hub* e *autoridade*
 - Atualizar scores de autoridade
Para cada nó *n*, o novo score de autoridade de *n* será a soma de todos os scores de *hub* das páginas que apontam para *n*
 - Atualizar scores de *hub*
Para cada nó *n*, o novo score de hub de *n* será a soma de todos os scores de *autoridade* das páginas que apontam para *n*
 - Normalizar scores de *hub* e *autoridade*

[Exercício Análise de Links – KNIME]





Obrigado!

Agente Educacional

Marcelo Pita

marcelo.pita@serpro.gov.br | #81 8794

Demais agentes educacionais sobre o assunto:

Sérgio M. Dias | sergio.dias@serpro.gov.br | #31 6539

Gustavo Torres | gustavo.gamatorres@serpro.gov.br | #31 6950