



# Neo4j: Predicción de enlaces

**Big Data Aplicado**

Julio Alberto López Gómez  
JulioAlberto.Lopez@uclm.es



**Big\_Data**  
Aplicado



Curso Especialización  
Inteligencia\_Artificial y  
Big\_Data

# Predictión de enlaces:

## Concepto:

- Los grafos son modelos de datos dinámicos
  - Evolucionan a lo largo del tiempo
  - Aparecen nuevos nodos y se crean/eliminan conexiones
- Predecir qué enlaces pueden formarse en un grafo puede ayudar:
  - Prevenir eventualidades
  - Adelantarse a los acontecimientos
  - ...

# Predictión de enlaces:

## Concepto:

- Existen multitud de métodos
- Basados en medidas de cercanía y de centralidad
  - Los nuevos enlaces tenderán a aparecer desde/sobre los nodos más relevantes
- Algunos métodos representativos:
  - Vecinos comunes
  - Adhesión preferente
  - Asignación de recursos
  - ...

# Predicción de enlaces:

Ejemplo:

- Creación de un grafo.

CREATE

```
(zhen:Person {name: 'Zhen'}) ,  
(praveena:Person {name: 'Praveena'}) ,  
(michael:Person {name: 'Michael'}) ,  
(arya:Person {name: 'Arya'}) ,  
(karin:Person {name: 'Karin'}) ,  
  
(zhen) - [:FRIENDS] -> (arya) ,  
(zhen) - [:FRIENDS] -> (praveena) ,  
(praveena) - [:WORKS WITH] -> (karin) ,  
(praveena) - [:FRIENDS] -> (michael) ,  
(michael) - [:WORKS WITH] -> (karin) ,  
(arya) - [:FRIENDS] -> (karin)
```

# Predictión de enlaces:

Ejemplo:

- Creación de un grafo.



# Predicción de enlaces:

## Vecinos comunes:

- Idea principal:
  - Dos extraños con un amigo en común tienen más posibilidad de conectarse entre sí que quienes no
- $CN(x, y) = |N(x) \cap N(y)|$ 
  - $N(x)$  es el conjunto de nodos adyacentes a  $x$
  - $N(y)$  es el conjunto de nodos adyacentes a  $y$

# Predictión de enlaces:

Vecinos comunes:

- Ejecución
  - Devuelve el número de vecinos comunes de Michael y Karin

```
MATCH (p1:Person {name: 'Michael'})  
MATCH (p2:Person {name: 'Karin'})  
RETURN  
gds.alpha.linkprediction.commonNeighbors(p1, p2) AS score
```



- Resultado: 1

# Predicción de enlaces:

## Adhesión preferencial:

- Idea principal:
  - Cuanto más conectado está un notado, es más probable que reciba nuevos enlaces.
- $PA(x, y) = |N(x)| \cdot |N(y)|$ 
  - $N(x)$  conjunto de nodos adyacentes a  $x$
  - $N(y)$  conjunto de nodos adyacentes a  $y$

# Predictión de enlaces:

## Adhesión preferencial:

- Ejecución
  - Devuelve el valor de adhesión preferencial para Michael y Karin

```
MATCH (p1:Person {name: 'Michael'})  
MATCH (p2:Person {name: 'Karin'})  
RETURN  
gds.alpha.linkprediction.preferential  
Attachment(p1, p2) AS score
```

- Resultado: 6

# Predicción de enlaces:

## Asignación de recursos:

- Métrica compleja que evalúa la cercanía entre nodos para determinar la posibilidad de que se produzca un enlace entre ellos.
- $RA(x, y) = \sum_{u \in N(x) \cap N(y)} \frac{1}{|N(u)|}$
- $N(u)$  es el conjunto de nodos adyacentes a  $u$

# Predictión de enlaces:

## Asignación de recursos:

- Ejecución
  - Devuelve el valor de asignación de recursos para Michael y Karin

```
MATCH (p1:Person {name: 'Michael'})  
MATCH (p2:Person {name: 'Karin'})  
RETURN  
gds.alpha.linkprediction.resourceAllocation(p1, p2) AS score
```

- Resultado: 0,333



# Neo4j: Predicción de enlaces

**Big Data Aplicado**

Julio Alberto López Gómez  
JulioAlberto.Lopez@uclm.es



**Big\_Data  
Aplicado**



Curso Especialización  
Inteligencia\_Artificial y  
Big\_Data