
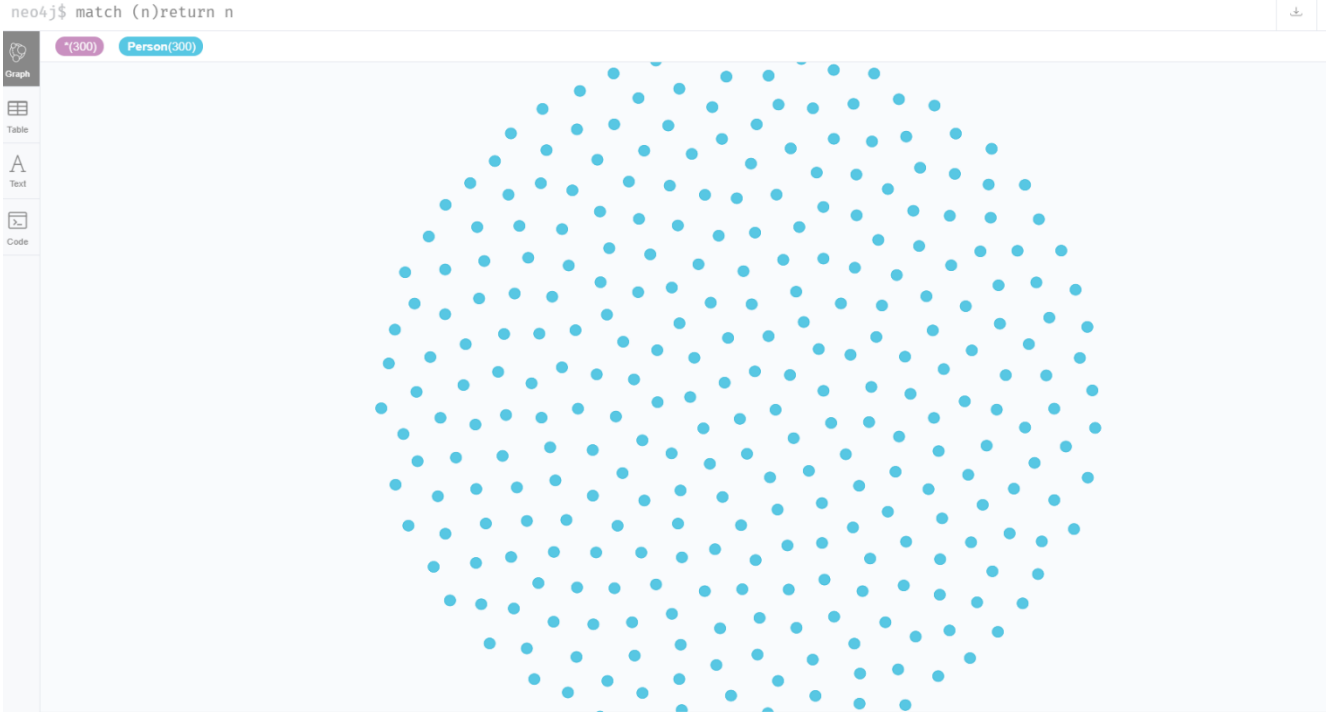
	VICERRECTORADO DOCENTE	Código: GUIA-PRL-001
	CONSEJO ACADÉMICO	Aprobación: 2016/04/06
Formato: Guía de Práctica de Laboratorio / Talleres / Centros de Simulación		

		FORMATO DE INFORME DE PRÁCTICA DE LABORATORIO / TALLERES / CENTROS DE SIMULACIÓN – PARA ESTUDIANTES	
CARRERA: Ingeniería en Sistemas		ASIGNATURA: Sistemas Expertos	
NRO. PRÁCTICA:	8	TÍTULO PRÁCTICA: KNN Clasificación de miembros del congreso utilizando algoritmos de similitud en Neo4j	
OBJETIVO ALCANZADO: <ul style="list-style-type: none"> - Adquirir destrezas en el manejo de algoritmos desarrollados por neo4j. - Entender la lógica de implementación en inteligencia artificial de los algoritmos de neo4j 			
<p align="center">ACTIVIDADES DESARROLLADAS</p> <p>1. La información sobre el conjunto de datos y el conjunto de datos en sí, exportamos los datos.</p> <p>LOAD CSV FROM "http://archive.ics.uci.edu/ml/machine-learning-databases/voting-records/house-votes-84.data" as row CREATE (p:Person) SET p.class = row[0], p.features = row[1..]</p> <div> <pre>neo4j\$ match (n) return n</pre>  </div>			

2. Verificamos la cantidad de votos.

MATCH (n:Person)
WHERE "?" in n.features
RETURN count(n)

neo4j\$ MATCH (n:Person) WHERE "?" in n.features RETURN co...

count(n)
203


3. Distribucion de votos faltantes

MATCH (p:Person)
WHERE '?' in p.features
WITH p,apoc.coll.occurrences(p.features,'?') as missing
RETURN missing,count(*) as times ORDER BY missing ASC

neo4j\$ MATCH (p:Person) WHERE '?' in p.features WITH p,ap...

missing	times
1	124
2	43
3	16
4	6
5	5
6	4
7	1
9	1
14	1


Started streaming 11 records in less than 1 ms and completed after 5 ms.

	VICERRECTORADO DOCENTE	Código: GUIA-PRL-001
	CONSEJO ACADÉMICO	Aprobación: 2016/04/06
Formato: Guía de Práctica de Laboratorio / Talleres / Centros de Simulación		

4. Eliminar nodos que no tienen votos

```
MATCH (p:Person)
WITH p,apoc.coll.occurrences(p.features,'?') as missing
WHERE missing > 6
DELETE p
```

neo4j\$ MATCH (p:Person) WITH p,apoc.coll.occu...

 Deleted 5 nodes, completed after 39 ms.

5. Dividimos el conjunto

```
MATCH (p:Person)
WITH p LIMIT 344
SET p:Training;
```

neo4j\$ MATCH (p:Person) WITH p LIMIT 344 SET ...

 Added 344 labels, completed after 29 ms.




```
MATCH (p:Person)
WITH p SKIP 344
SET p:Test;
```

neo4j\$ MATCH (p:Person) WITH p SKIP 344 SET p...

 Added 86 labels, completed after 12 ms.



	VICERRECTORADO DOCENTE	Código: GUIA-PRL-001
	CONSEJO ACADÉMICO	Aprobación: 2016/04/06
Formato: Guía de Práctica de Laboratorio / Talleres / Centros de Simulación		

6. Transformar a vector de características

MATCH (n:Person)

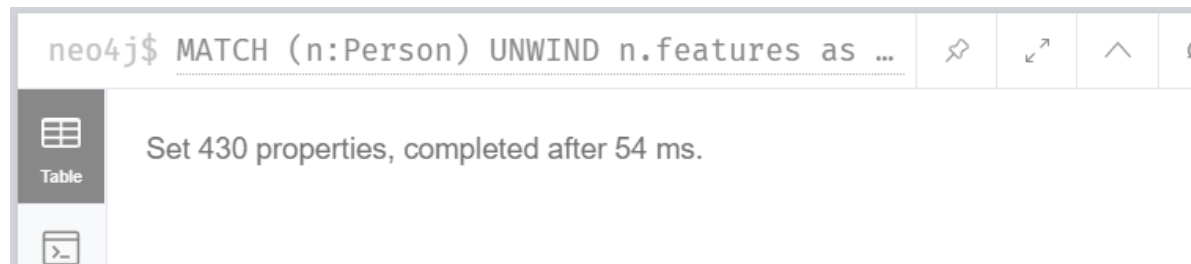
UNWIND n.features as feature

WITH n,collect(CASE feature WHEN 'y' THEN 1

WHEN 'n' THEN 0

ELSE 0.5 END) as feature_vector

SET n.feature_vector = feature_vector



neo4j\$ MATCH (n:Person) UNWIND n.features as ...

Set 430 properties, completed after 54 ms.

7. Algoritmo KNN

MATCH (test:Test)

WITH test,test.feature_vector as feature_vector

CALL apoc.cypher.run('MATCH (training:Training) WITH training, gds.alpha.similarity.euclideanDistance(\$feature_vector, training.feature_vector) AS similarity

ORDER BY similarity ASC LIMIT 3

RETURN collect(training.class) as classes',

{feature_vector:feature_vector}) YIELD value WITH test.class as class,

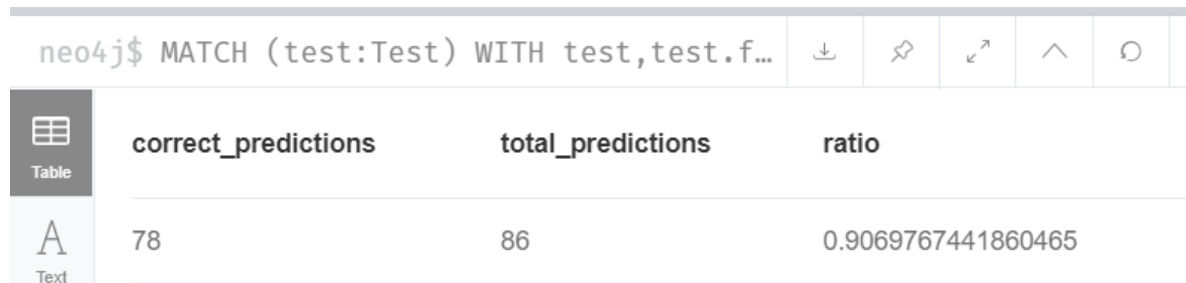
apoc.coll.sortMaps(apoc.coll.frequencies(value.classes),

'^count')[-1].item as predicted_class

WITH sum(CASE when class = predicted_class THEN 1 ELSE 0 END) as correct_predictions, count(*) as

total_predictions RETURN correct_predictions,total_predictions,

correct_predictions / toFloat(total_predictions) as ratio



neo4j\$ MATCH (test:Test) WITH test,test.f...

correct_predictions	total_predictions	ratio
78	86	0.9069767441860465

Nombre de los estudiantes: Fanny Gutama