

Neo4j-Pathway Embeddings

Nov. 2nd 2020

Loading data into Database

It's just the Pathways that I already had downloaded...

Neo4j 4.1.3

5,434 nodes (2 labels)

15,672 relationships (1 types)

**MATCH (n) RETURN labels(n) AS NodeType,
count(n) AS NumberOfNodes;**

NodeType	NumberOfNodes	
1	["Gene"]	4694
2	["Pathway"]	740

MATCH (p:Pathway)

RETURN p.Origin AS Source, count(p.Origin) AS Count

ORDER BY Count ASC;

Source	Count	
1	"Kegg"	200
2	"Biocarta"	254
3	"HumanCyc"	286

```
$ CALL db.schema.visualization()
```



Graph



Table



Text

*(2)

Gene(1)

Pathway(1)

*(1)

BELONGS_TO(1)



```
CALL gds.alpha.node2vec.stream({
  nodeProjection: "Pathway",
  relationshipProjection: {
    eroad: {
      type: "BELONGS_TO",
      orientation: "UNDIRECTED"
    }
  },
  embeddingSize: 6,
  iterations: 10,
  walkLength: 10
})
YIELD nodeId, embedding
RETURN gds.util.asNode(nodeId).Name AS pathway, embedding
LIMIT 5;
```

```
neo4j$ CALL gds.alpha.node2vec.stream({ nodeProjection: "Pathway", relationshipProjection: { e...
```

	pathway	embedding
1	"antisense pathway"	[-0.5685299634933472, -0.1976551115512848, 0.10311362147331238, 0.20070020854473114, -0.6220663189888, -0.9396705031394958]
2	"hemoglobins chaperone"	[0.10599073767662048, 0.5979461073875427, 0.29990530014038086, 0.8381250500679016, -0.3769611418247223, 0.6305319666862488]
3	"srebp control of lipid synthesis"	[-0.7643737196922302, 0.8954192996025085, -0.7893279790878296, -0.8813167214393616, 0.14982885122299194, 0.6395041346549988]
4	"multi-drug resistance factors"	[-0.7179145216941833, -0.3299995958805084, -0.08277614414691925, 0.7398416996002197, -0.7939758896827698, -0.08961626142263412]
5	"inhibition of matrix metalloproteinases"	[0.4254385828971863, 0.7173682451248169, 0.07536022365093231, -0.7533583045005798, -0.5216968059539795, 0.041713450103998184]

Started streaming 5 records after 30 ms and completed after 8584 ms.





