



**UNIVERSIDADE FEDERAL DE SÃO PAULO  
INSTITUTO DE CIÊNCIA E TECNOLOGIA  
BACHARELADO EM CIÊNCIA E TECNOLOGIA**

**Projeto 03:**

***Rede GNG - Growing Neural Gas***

**Nome:** Willian Dihanster Gomes de Oliveira **RA:** 112269

**SÃO JOSÉ DOS CAMPOS  
2018**

### Projeto 3 – Experimentos GNG

Implementar ou utilizar alguma biblioteca pronta.

Utilizar o algoritmo como ferramenta de agrupamento em pelo menos 3 datasets.

Analise o impacto dos seguintes parâmetros:

- Vida das ligações (arestas).
- Intervalo de inserção de novos neurônios.

#### Implementação

Para a realização deste trabalho, foi usado o código-fonte disponibilizado por Adrien Guille em seu *github*, feito em *python*.

Os principais parâmetros foram setados inicialmente por:

$a_{max} = 10$  (idade máxima das arestas)

$l = 100$  (a cada  $l$  iterações insere novos neurônios)

#### Conjunto de Dados 1: Moons

Para a base de dados *Moons* e o conjunto de parâmetros iniciais do algoritmo encontrado, temos os seguintes resultados das Figuras 1 e 2.

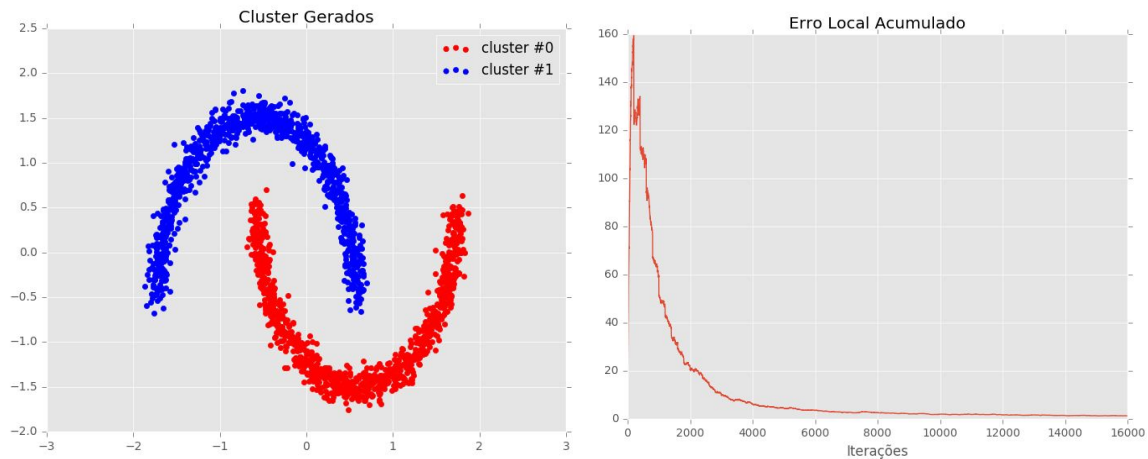


Figura 1: Dois clusters gerados e gráfico do erro local acumulado para a rede Moons.

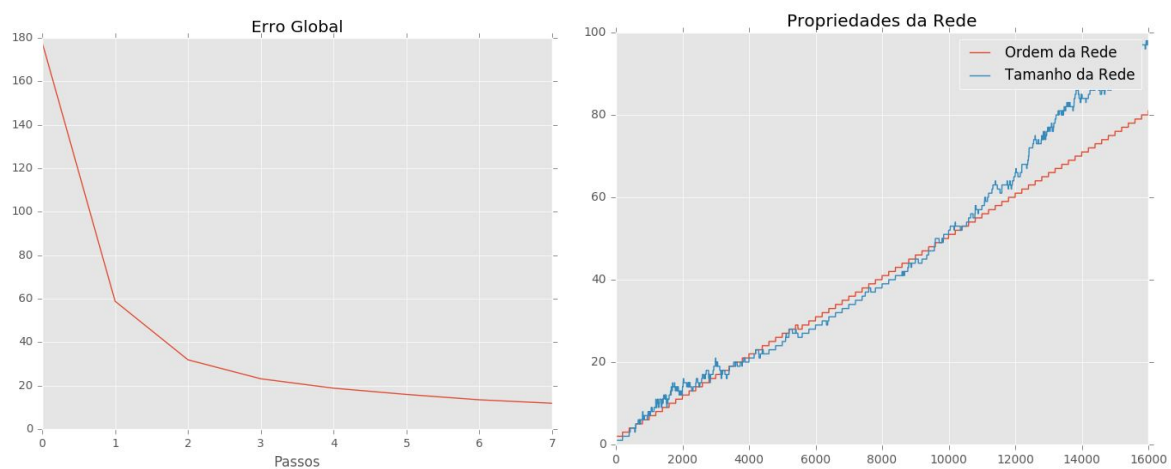


Figura 2: Gráfico do erro global e dados de propriedades da rede.

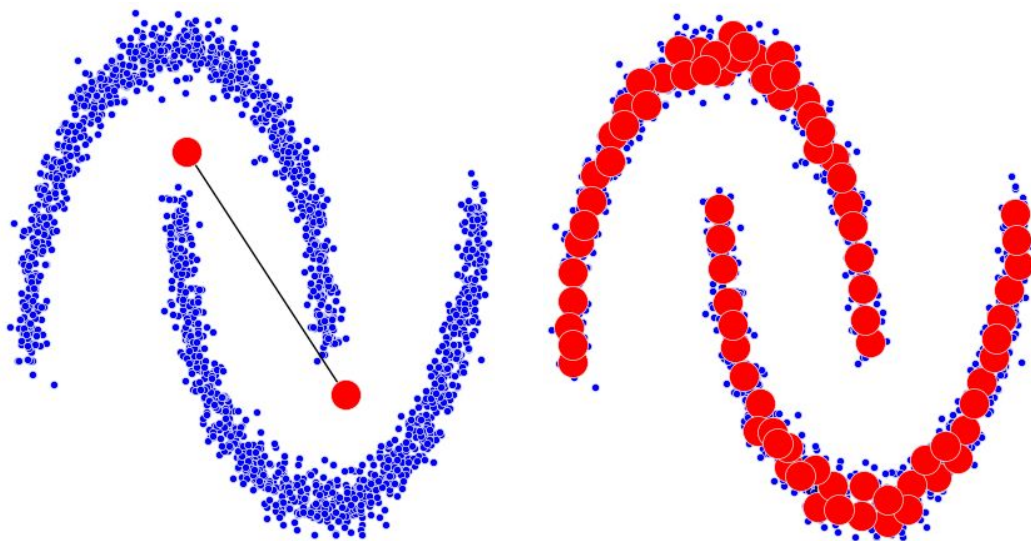


Figura 3: Rede gerada na 1ª iteração e na última iteração do algoritmo.

## Dataset 2: Circles

Com o conjunto de parâmetros iniciais do algoritmo encontrado, temos os seguintes resultados das Figuras 4 e 5.

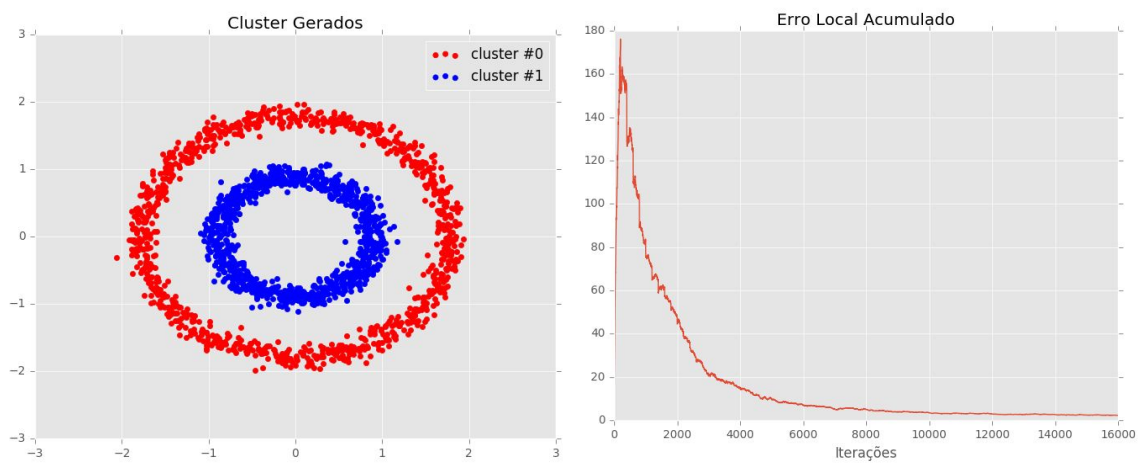


Figura 4: Dois clusters gerados e gráfico do erro local acumulado para a rede *Circles*.

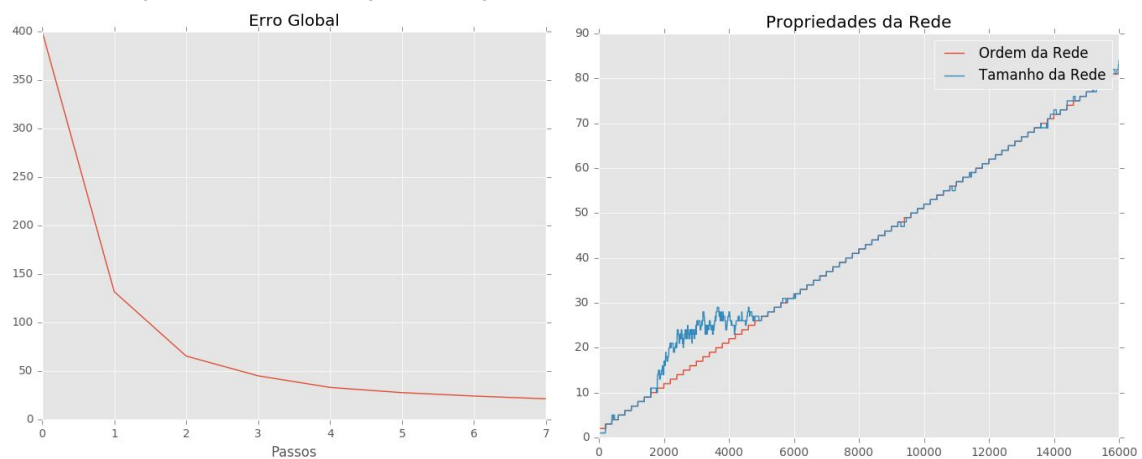


Figura 5: Gráfico do erro global e dados de propriedades da rede.

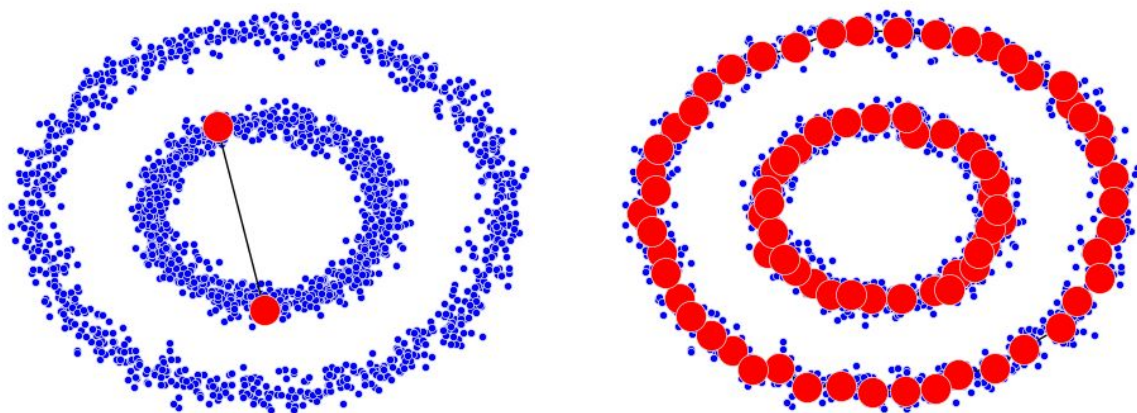


Figura 6: Rede gerada na 1ª iteração e na última iteração do algoritmo.

### Dataset 3: Blobs

Com o conjunto de parâmetros iniciais do algoritmo encontrado, temos os seguintes resultados das Figuras 7 e 8.

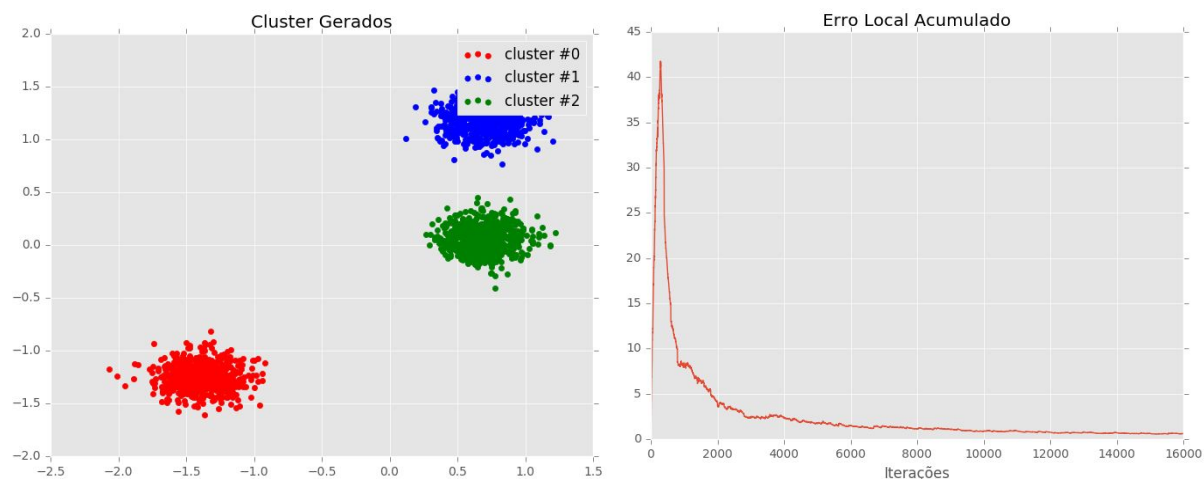


Figura 7: Dois clusters gerados e gráfico do erro local acumulado para a rede *Blobs*.



Figura 8: Gráfico do erro global e dados de propriedades da rede.

Também é possível observar a evolução da rede conforme o número de iterações. Para os parâmetros iniciais, temos o seguinte resultado da Figura 9.

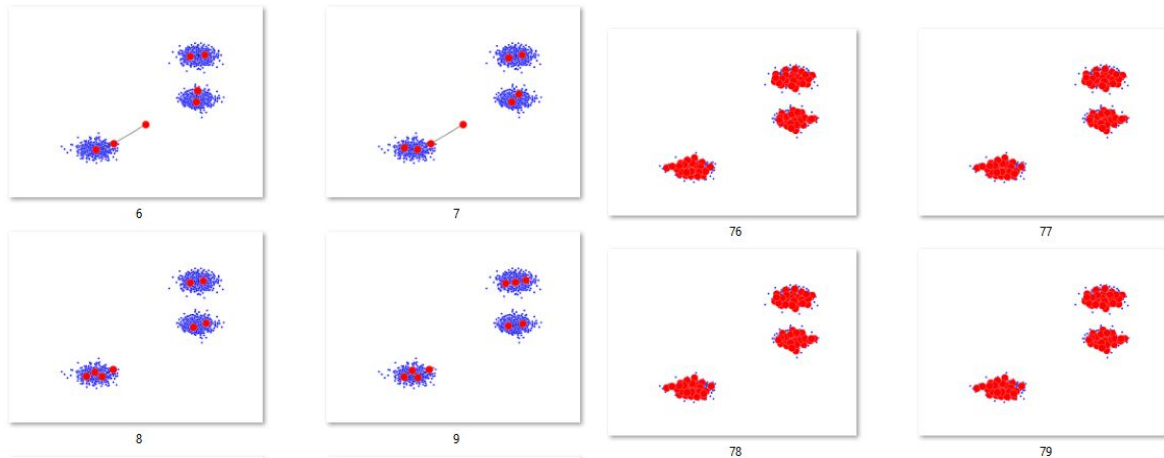


Figura 9: Rede gerada conforme o número de iterações com os parâmetros padrões do algoritmo.

Pode-se observar que a partir da iteração 8, já não existe uma aresta “solta”. No final, podemos ver que a rede encontrou o resultado esperado.

Esse rede considera como a idade máxima das aresta,  $a_{max} = 10$ . Sendo assim, novos experimentos foram realizados, considerando novos valores para  $a_{max}$ .

Para  $a_{max} = 5$ , temos os seguintes resultados da Figura 10.

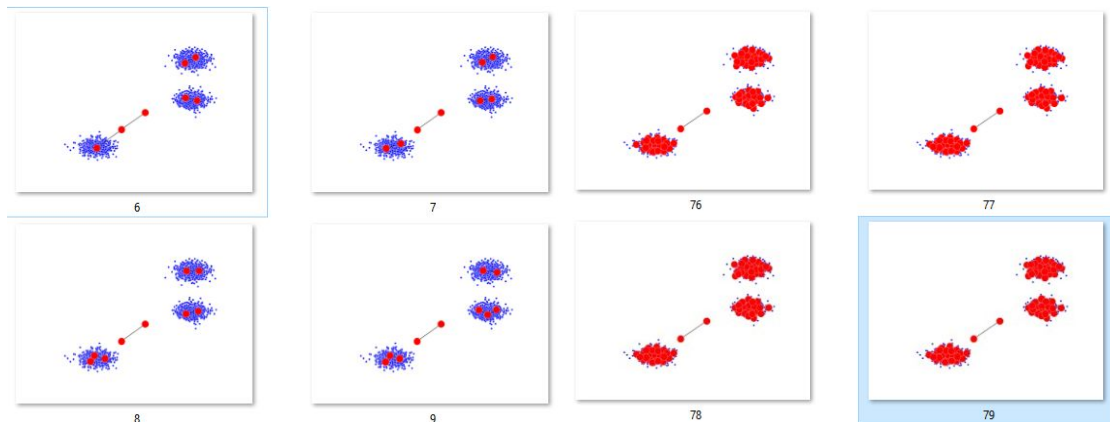


Figura 10: Rede gerada conforme o número de iterações com  $a_{max} = 5$ .

Nota-se que mesmo no final das iterações, a rede ainda contém uma aresta “solta” entre os clusters.

Para  $a_{max} = 10$ , temos os seguintes resultados da Figura 11.

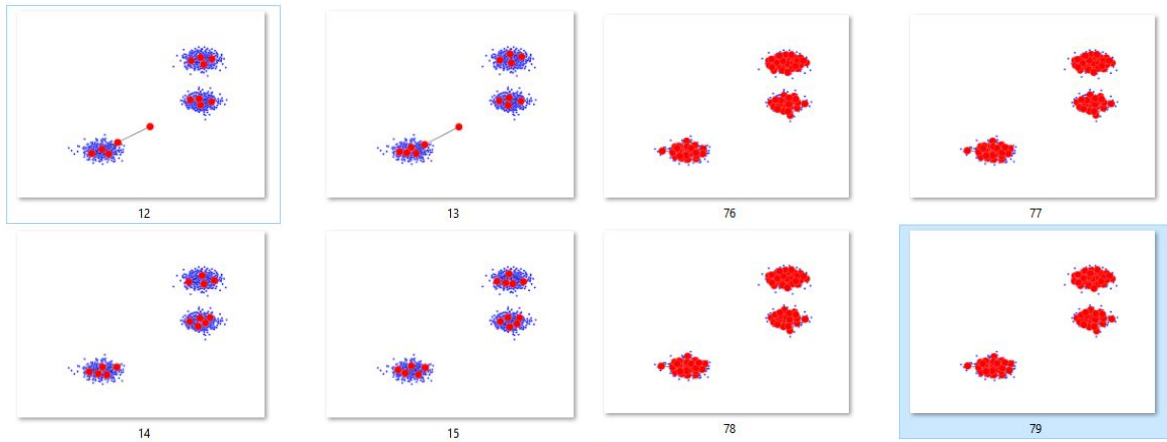


Figura 11: Rede gerada conforme o número de iterações com  $a_{max} = 20$ .

Com o dobro do valor de  $a_{max}$  padrão da rede gerada, nota-se que houve um maior número de iterações para a aresta “solta” sumir, em relação ao valor de  $a_{max}$  padrão. No entanto, o resultado final, é praticamente o mesmo para  $a_{max} = 10$ .

Diminuindo o tempo de geração de arestas pela metade, temos o seguinte resultado da Figura 12.

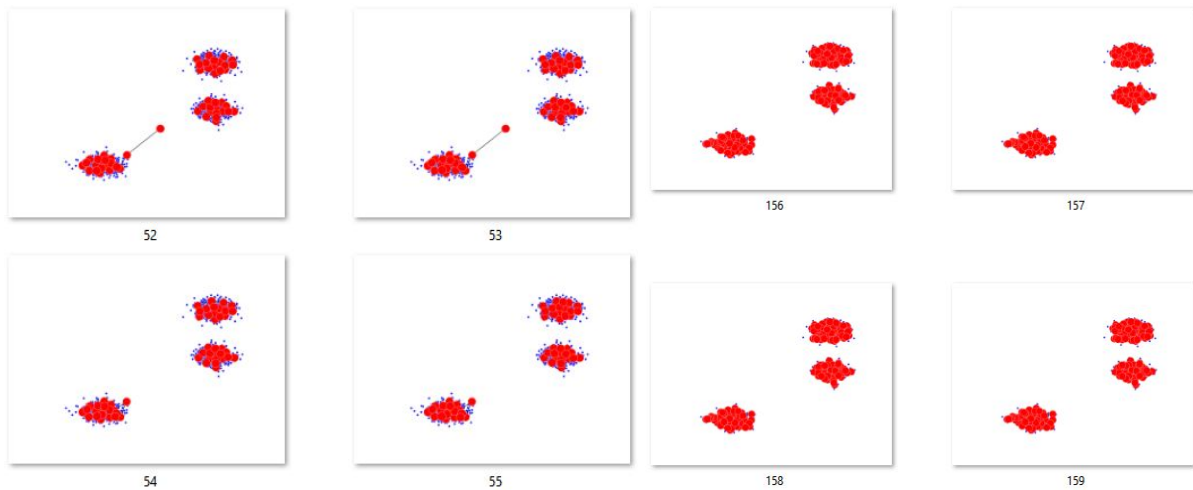


Figura 12: Rede gerada conforme o número de iterações com  $l = 100$ .

Pode-se notar que na 53-*ésima* iteração, ainda havia a existência de uma aresta “solta”. Além disso, o algoritmo leva mais iterações para a rede chegar em sua forma final.

Para  $l = 400$ , temos o resultado da Figura 13.

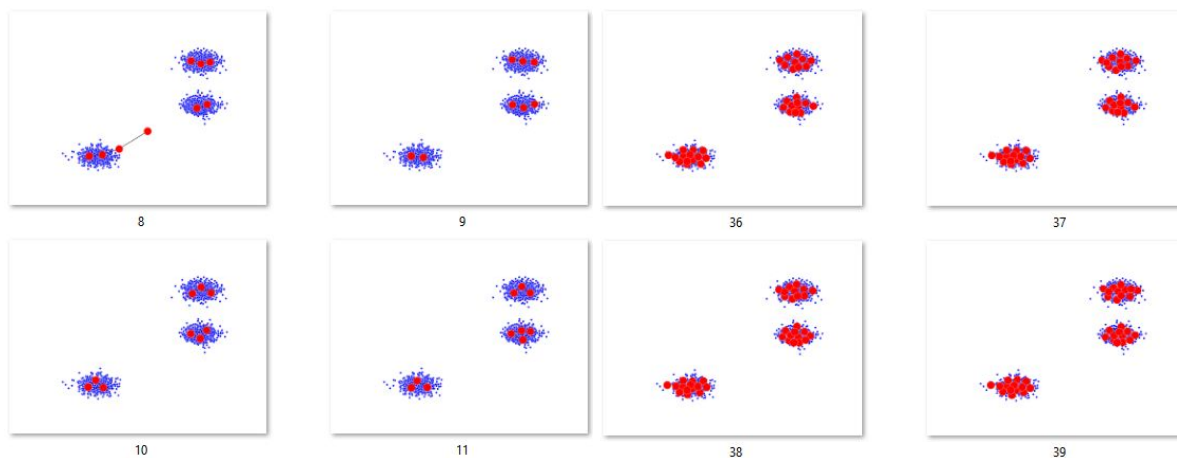


Figura 13: Rede gerada conforme o número de iterações com  $l = 400$ .

Nota-se uma pequena diferença em relação ao resultado original. Onde na original, a aresta “solta” sumiu na 7ª iteração, enquanto que nessa configuração, sumiu na 8ª iteração. No entanto, chegou no mesmo resultado com menos iterações (40 contra 80).

### Conclusões

Com tais experimentos pode-se concluir a eficácia do algoritmo na GNG na representação de dados, sobretudo, dados de agrupamento. Pois conseguiu representar bem os dados originais. Mas deve-se atentar em alguns parâmetros que podem fazer com que a rede demore mais a convergir ou não consiga representar os dados de forma razoável.