



Self Organized Maps (Kohonen)
Profa. Dra. Roseli Aparecida Francelin Romero
Disciplina: SCC 0270 - Redes Neurais

Lucas Yudi Sugi - 9293251

Introdução

As Self-Organized-Maps (SOM) são redes neurais auto organizáveis da classe dos algoritmos não supervisionados. A primeira rede SOM existente foi criada por Teuvo Kohonen e tem como finalidade principal transformar dados de alta dimensão para uma menor de maneira que seja possível a visualização. A rede Kohonen consegue tal feito por meio de agrupamentos realizados nos dados.

Objetivo

Implementar a rede Kohonen e testá-la no dataset wine com o objetivo de verificar os agrupamentos finais de cada classe. Link do dataset utilizado:

<https://archive.ics.uci.edu/ml/datasets/wine>

Implementação

A rede Kohonen foi implementada por meio de uma classe na qual a sua instância necessita dos seguintes parâmetros:

- Data: Dados que serão agrupados.
- Grid: Grid dos neurônios da rede.
- Alpha: Taxa de aprendizado.
- Sigma: Raio de alcance do neurônio ganhador.

Após a criação do objeto deve-se chamar a função train passando quantas épocas o algoritmo irá executar. Internamente a função realiza o treinamento da rede clássica de kohonen. Basicamente os seguintes passos são realizados (pseudocódigo):

1. Cálculo da constante de tempo.
2. Para cada época:
 - 2.1. Cálculo de alpha e sigma.
 - 2.1.1. Para cada amostra no dataset:
 - 2.1.1.1. Cálculo da matriz de distâncias.
 - 2.1.1.2. Cálculo do neurônio vencedor(Maior ativação).
 - 2.1.1.2.1. Para cada neurônio no grid:
 - 2.1.1.2.1.1. Distância topográfica entre neurônio vencedor e o atual.

- 2.1.1.2.1.2. Cálculo da influência do neurônio vencedor.
- 2.1.1.2.1.3. Atualização dos pesos.

Tal pseudocódigo pode ser visualizado de maneira mais detalhada no arquivo `kohonen.py`, onde se encontra a implementação em python da rede.

Teste

O SOM criado para o dataset wine possui um grid 10x10 com taxa de aprendizado (alpha) de 0.5 e foi executado por 100 épocas. Após o seu treinamento foi gerado um gráfico contendo os agrupamentos dos neurônios.

É importante salientar que a visualização utilizada neste trabalho é diferente das usuais, i.e, não foi usado a Unified Distance Matrix (U-Matrix) nem os heatmap's de cada atributo. Nossa visualização foi criada de maneira muito mais modesta: Para cada neurônio no grid é verificado qual amostra é mais verossímil, i.e, possui mais chances de ativar o nó.

Conhecendo tal amostra, sua classe é atribuída para o neurônio. Assim, é possível identificar o rótulo de cada nó. Assim, tal método de visualização é muito simples e não está totalmente correto, pois nem sempre todos os neurônios serão vencedores. Isso significa que há nodos sobre influência de várias classes. Logo, é possível que haja ruído no mapa.

Abaixo há o gráfico gerado:

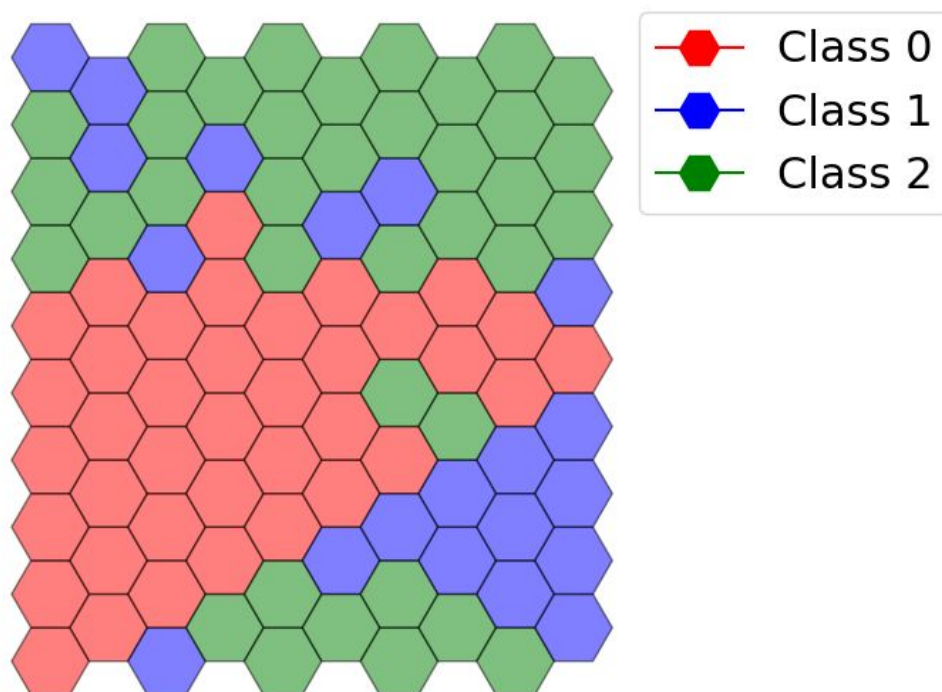


Figura 1: Mapa dos neurônios da rede separados pelas classes do wine.

Podemos ver pela figura que há uma boa separação nos dados, e que o ruído causado pode provir tanto da forma como é mostrado os dados quanto da sobreposição natural entre as classes. Para o nosso problema acreditamos que o fato de haver algum espalhamento das cores ocorre por causa da sobreposição natural das classes.

Conclusão

Na seção de teste fica evidente pela figura que o mapa de Kohonen cumpre o que promete, i.e, realiza o agrupamento dos dados. Podemos verificar que as três classes do dataset wine foram separadas em três grupos que possuem cortes diagonais no espaço bidimensional.

Assim, a rede foi capaz de permitir a visualização de um dataset com dimensão 13 para a dimensão 2.

Portanto, Kohonen é um algoritmo útil quando se tem dados de alta dimensão e deseja-se realizar a visualização dos dados.

Referências

<http://www.computacaointeligente.com.br/algoritmos/mapas-auto-organizaveis-som/>