

Relatório Exercício 6

Aluno: João Guilherme Madeira Araújo

Número USP: 9725165

Relatório do Exercício 6:

O exercício pediu a implementação de um método de redução de dimensionalidade e análise de dados chamado Self Organizing Map, SOM, ou mapa de Kohonen, no dataset wine. Devemos mostrar como SOM nos ajuda a melhor visualizar, entender e agrupar o dataset.

Datasets

Foi usado o dataset wine, que contém 178 vinhos em 3 classes distintas, cada vinho tem 13 características como teor alcoólico, percentual de ácido málico, ash e outros.

O Algoritmo

O código-fonte do Self Organizing Map está no arquivo SOM.py. O algoritmo tem uma grid com vários vetores, em cada passe para cada vetor de features da entrada nós computamos o vetor do mapa mais próximo da amostra, em seguida atualizamos esse vetor e seus vizinhos para ficarem mais próximos do feature vector da amostra. O mapa gerado tem dimensão 13 por 13, em cada nó temos representado com cores as classes dos inputs que esse nó é o best matching unit.

Conclusões

Podemos observar que o *Self Organizing Map* resultante (Figura 1), demonstra um bom agrupamento das 3 classes de vinho, porém podemos ver que as classes tem alguma mistura, representado pelos nós com cores mistas.

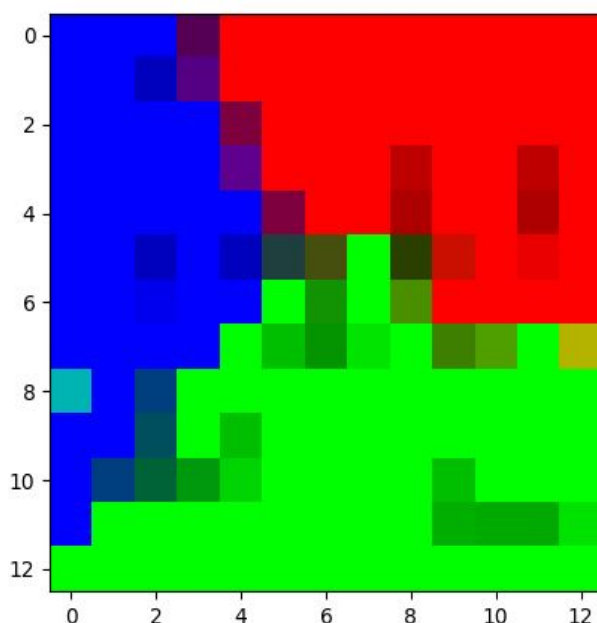


Figura 1: *Self Organizing Map* do dataset wine gerado pelo programa, cada classe é identificada por uma cor, um nó que é a BMU de vinhos de classes diferentes tem como cor a ponderada das cores das classes dos vinhos.