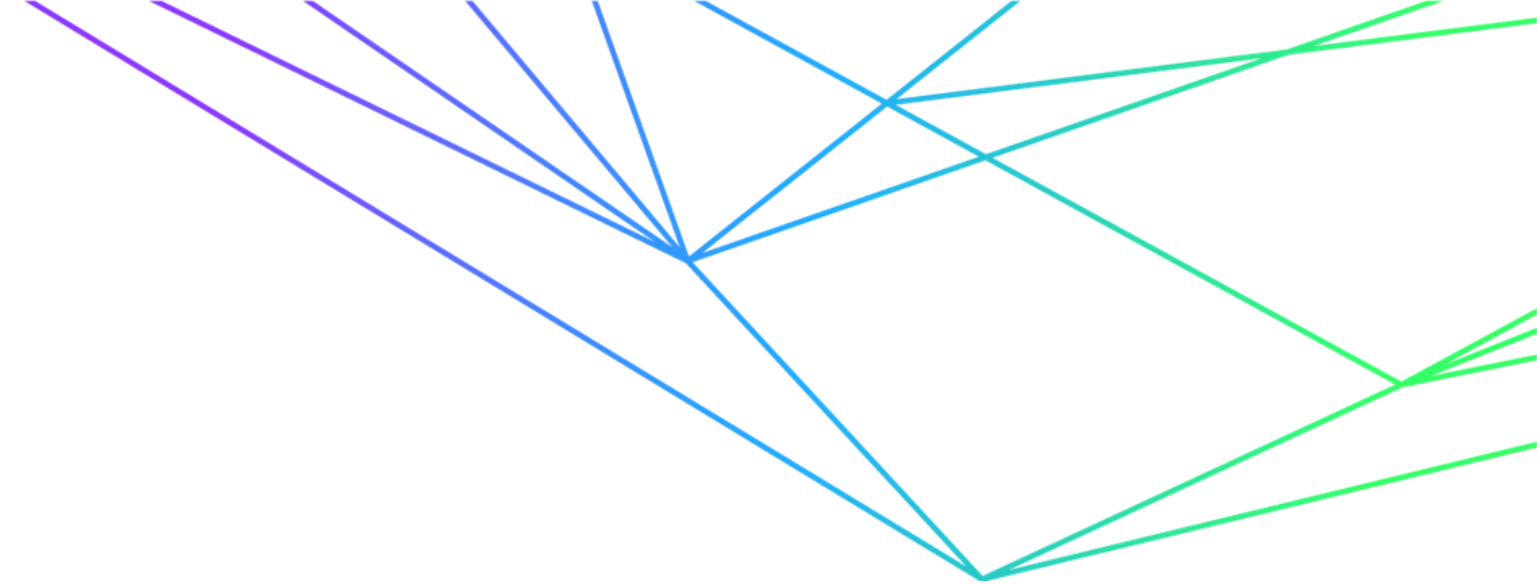


Clusterização

Agrupando objetos similares



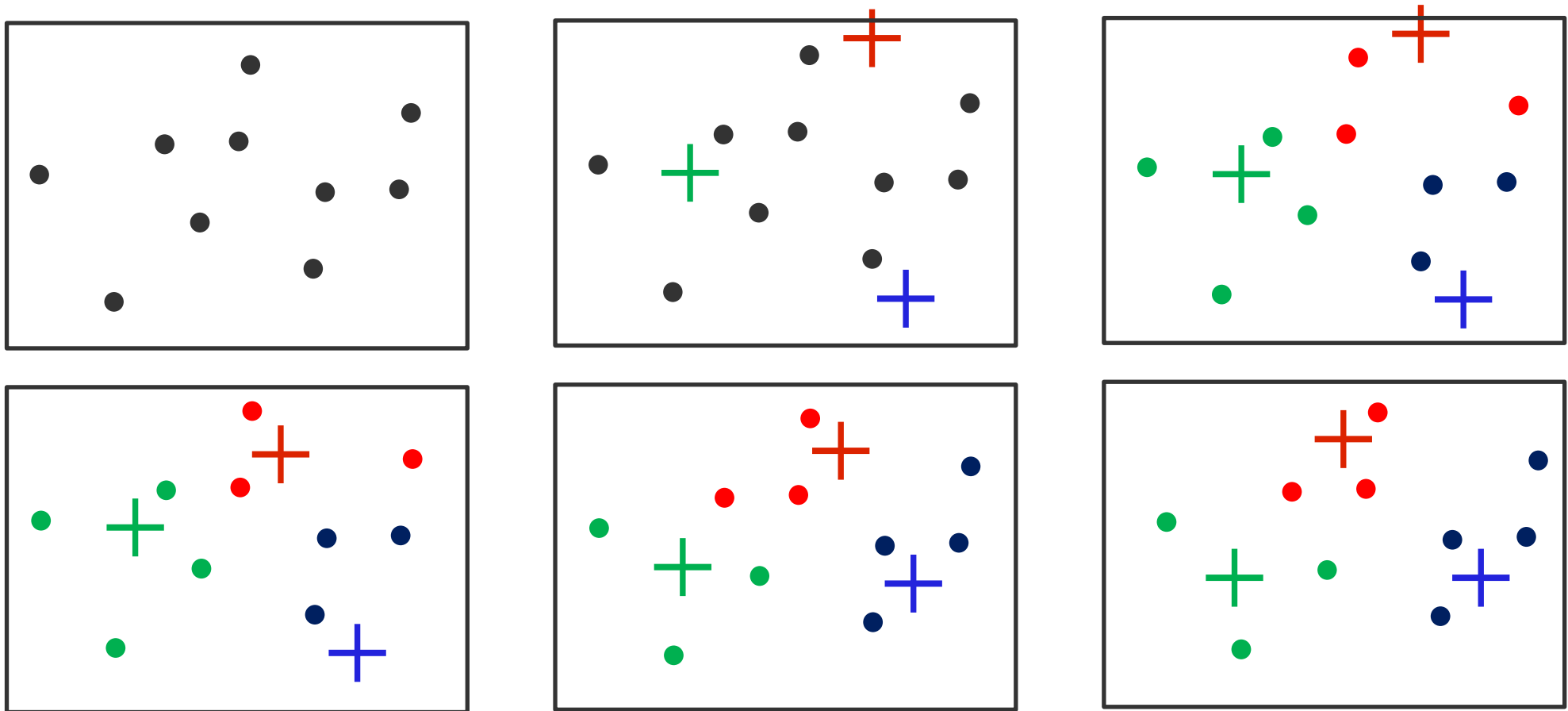
Classificação

Métodos

- K-Means
- Mean-Shift
- Density-Based Spatial Clustering of Applications with Noise (DBSCAN)
- Expectation–Maximization (EM) Clustering using Gaussian Mixture Models (GMM)
- Hierarchical Clustering

K-Means

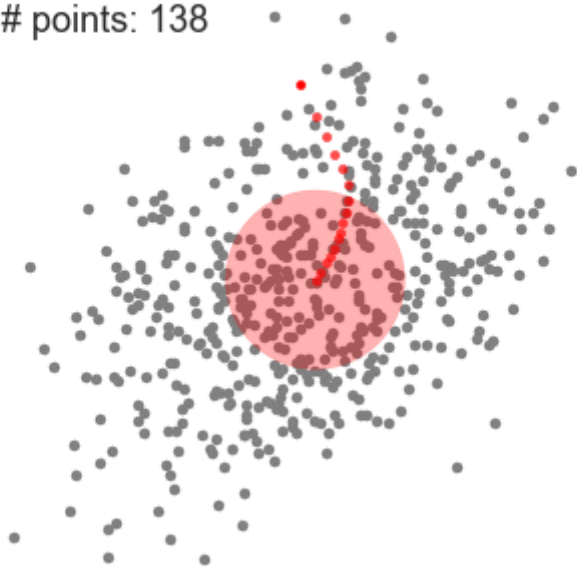
Se baseia na ideia de que o centro do cluster, chamado centroide, pode ser uma boa representação do cluster.



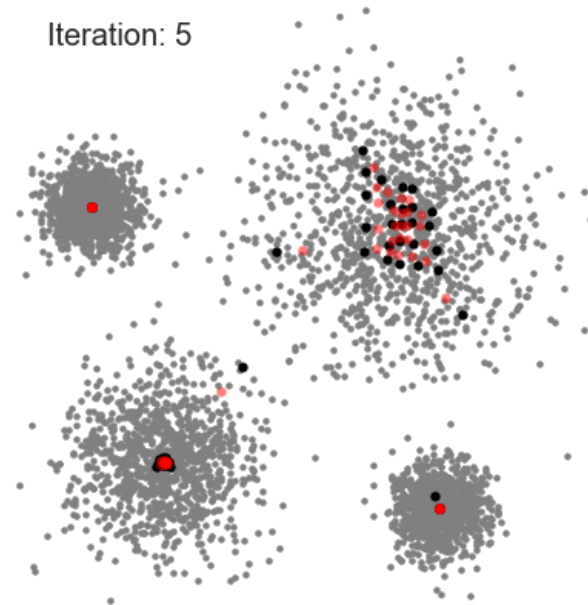
Mean-Shift

Move-se para a área de maior densidade, como o centro de gravidade das amostras, sem a necessidade de informar o número de grupos envolvidos

points: 138

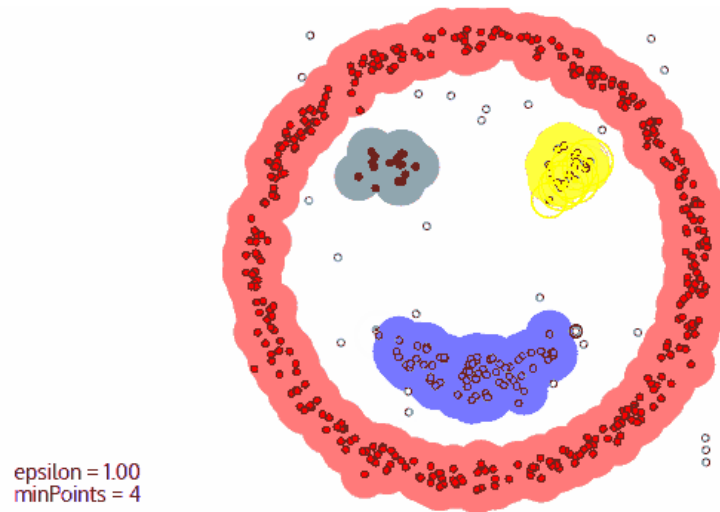


Iteration: 5



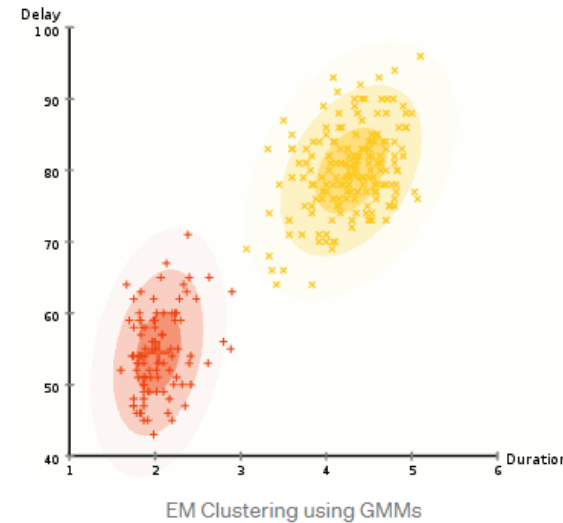
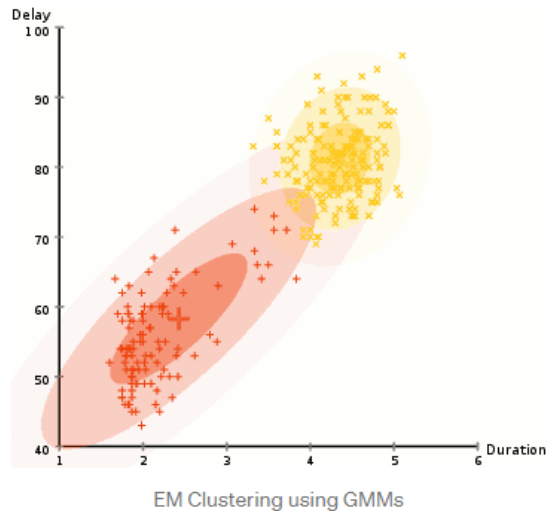
DBSCAN - Density-Based Spatial Clustering of Applications with Noise

Visita todos os pontos adicionando os que estão dentro de uma distância limite, se o número de pontos é menor que um limite estabelecido é considerado ruído.



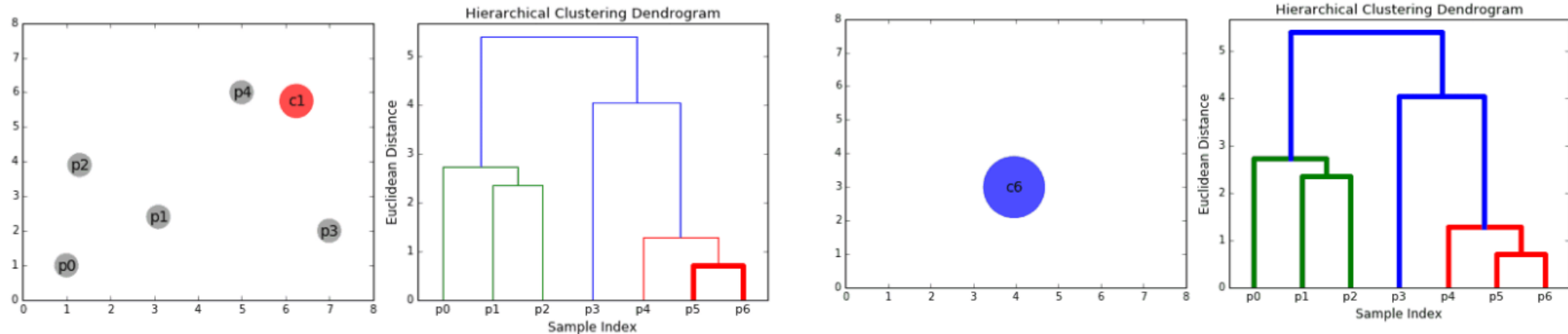
Expectation–Maximization (EM) Clustering using Gaussian Mixture Models (GMM)

Faz o ajuste de Gaussianas para descrever cada cluster.



Hierarchical Clustering

Clusterização Hierárquica pode ser tanto top-down or bottom-up. Na bottom-up cada ponto é considerado um cluster e a cada interação os clusters mais próximos são unificados até restar apenas um único cluster. O número final de cluster pode ser decido apenas parando o processo antes.





Semantix[®]

All about data

hader.azzini@semantix.com.br

www.semantix.com.br