



Gaussian Mixture Model - GMM

Aprendizado de Máquina

elias.rodriques@paulista.ifpe.edu.br



Aula passada: kMeans

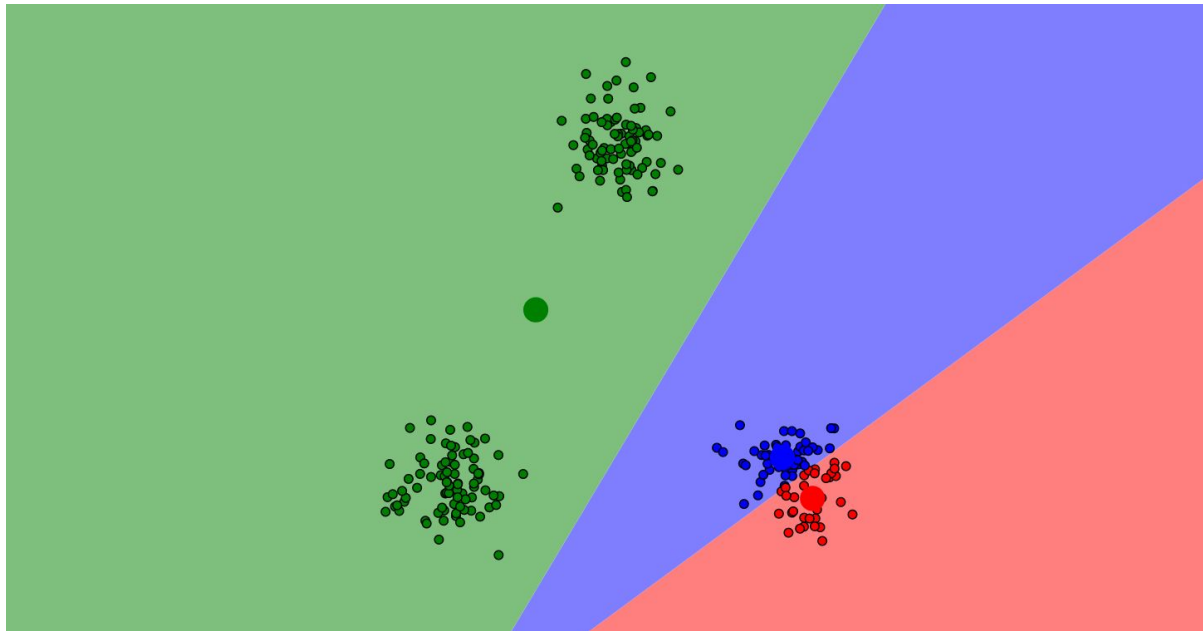
Prós:

- Simples de implementar
- Escalável.
- Garante convergência
- Facilmente adaptável a novos exemplos

Contras:

- Dependente da inicialização dos centróides
- Dificuldade com desbalanceamento nos aglomerados de dados
- Pressupõe aglomerados de dados esféricos
- Influenciado por outliers

Visualizing k-means clustering





Exemplo sintético

Dados sintéticos

- Padrões gerados aleatoriamente
 - 2D
 - 2 aglomerados de padrões
 - Seguindo distribuição Normal
- Função Densidade Probabilidade
 - Média e covariância parametrizados
- Classificador Bayesiano
 - Erro de 10,75%

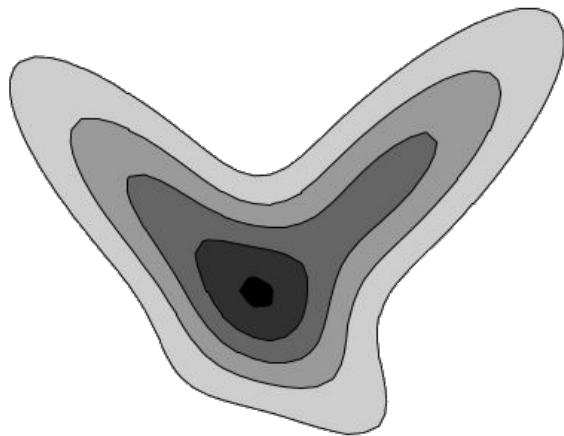
Agrupamento por KMeans

- KMeans
 - Recebe os padrões gerados aleatoriamente
 - Recebe $K=2$
 - Agrupa os padrões em torno de 2 centros
- Função Densidade Probabilidade
 - Média e covariância calculados para as amostras de cada grupo de padrões
- Classificador Bayesiano
 - Erro de 85,50%

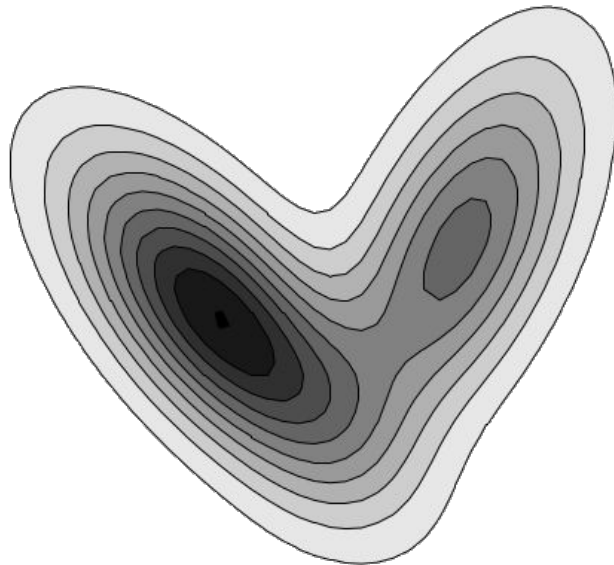


Exemplo sintético

Dados sintéticos

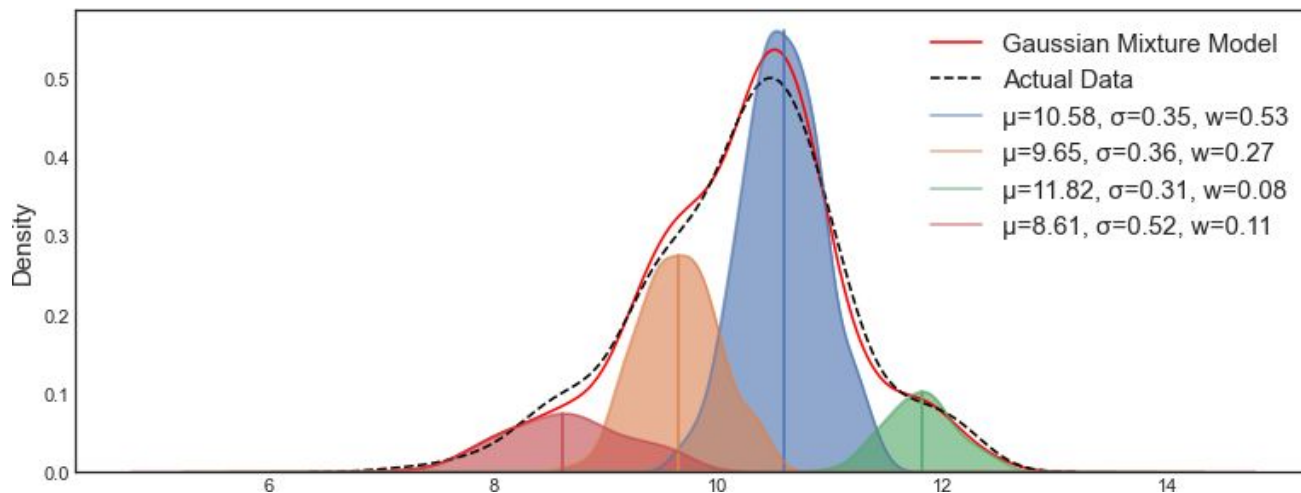


Agrupamento por KMeans



Gaussian Mixture Model (GMM)

- Técnica de agrupamento
- Busca pela melhor mistura de gaussianas para aproximar a geração dos dados





Exemplo sintético

Dados sintéticos

- Padrões gerados aleatoriamente
 - 2D
 - 2 aglomerados de padrões
 - Seguindo distribuição Normal
- Função Densidade Probabilidade
 - Média e covariância parametrizados
- Classificador Bayesiano
 - Erro de 10,75%

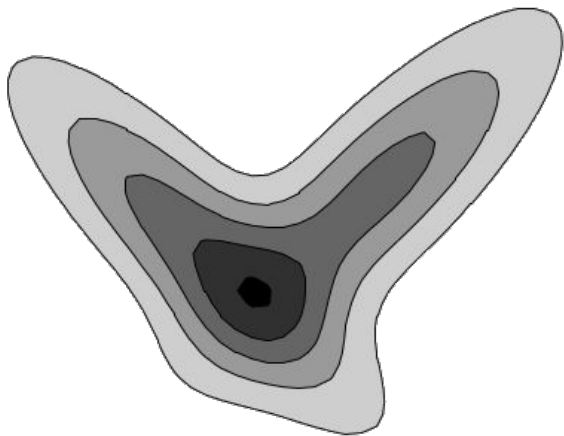
Agrupamento por GMM

- GMM
 - Recebe os padrões gerados aleatoriamente
 - Recebe $M=2$
 - Agrupa os padrões em torno de 2 centros
- Função Densidade Probabilidade
 - Média e covariância calculados para as amostras de cada grupo de padrões
- Classificador Bayesiano
 - Erro de 11,00%

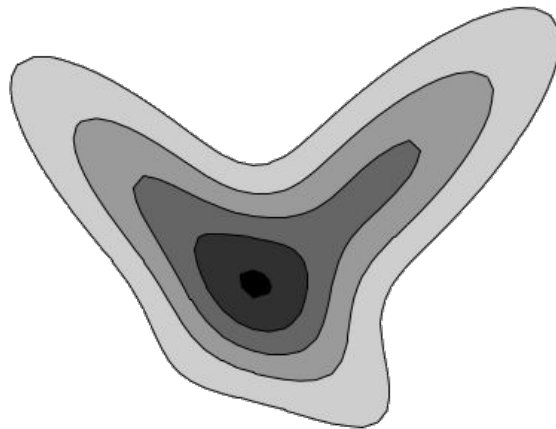


Exemplo sintético

Dados sintéticos



Agrupamento por GMM





Gaussian Mixture Model (GMM)

- Formalização matemática sólida
- Utiliza algoritmo de Expectation Maximization (EM)
 - Maximizar a verossimilhança do modelo

[A Gentle Tutorial of the EM Algorithm
and its Application to Parameter
Estimation for Gaussian Mixture and
Hidden Markov Models](#)



Gaussian Mixture Model (GMM)

- GMM está implementada no Scikit-Learn
 - [sklearn.mixture.GaussianMixture](https://scikit-learn.org/stable/modules/mixture.html)

```
>>> import numpy as np
>>> from sklearn.mixture import GaussianMixture
>>> X = np.array([[1, 2], [1, 4], [1, 0], [10, 2], [10, 4], [10, 0]])
>>> gm = GaussianMixture(n_components=2, random_state=0).fit(X)
>>> gm.means_
array([[10.,  2.],
       [ 1.,  2.]])
>>> gm.predict([[0, 0], [12, 3]])
array([1, 0])
```

>>>

Seg às 09:00

<https://meet.google.com/ngn-vjwh-qhd>

Dúvidas: elias.rodrigues@paulista.ifpe.edu.br

