



MACHINE LEARNING APLICADO À ANÁLISE DE DADOS

AULA 8 – 25/03/2021

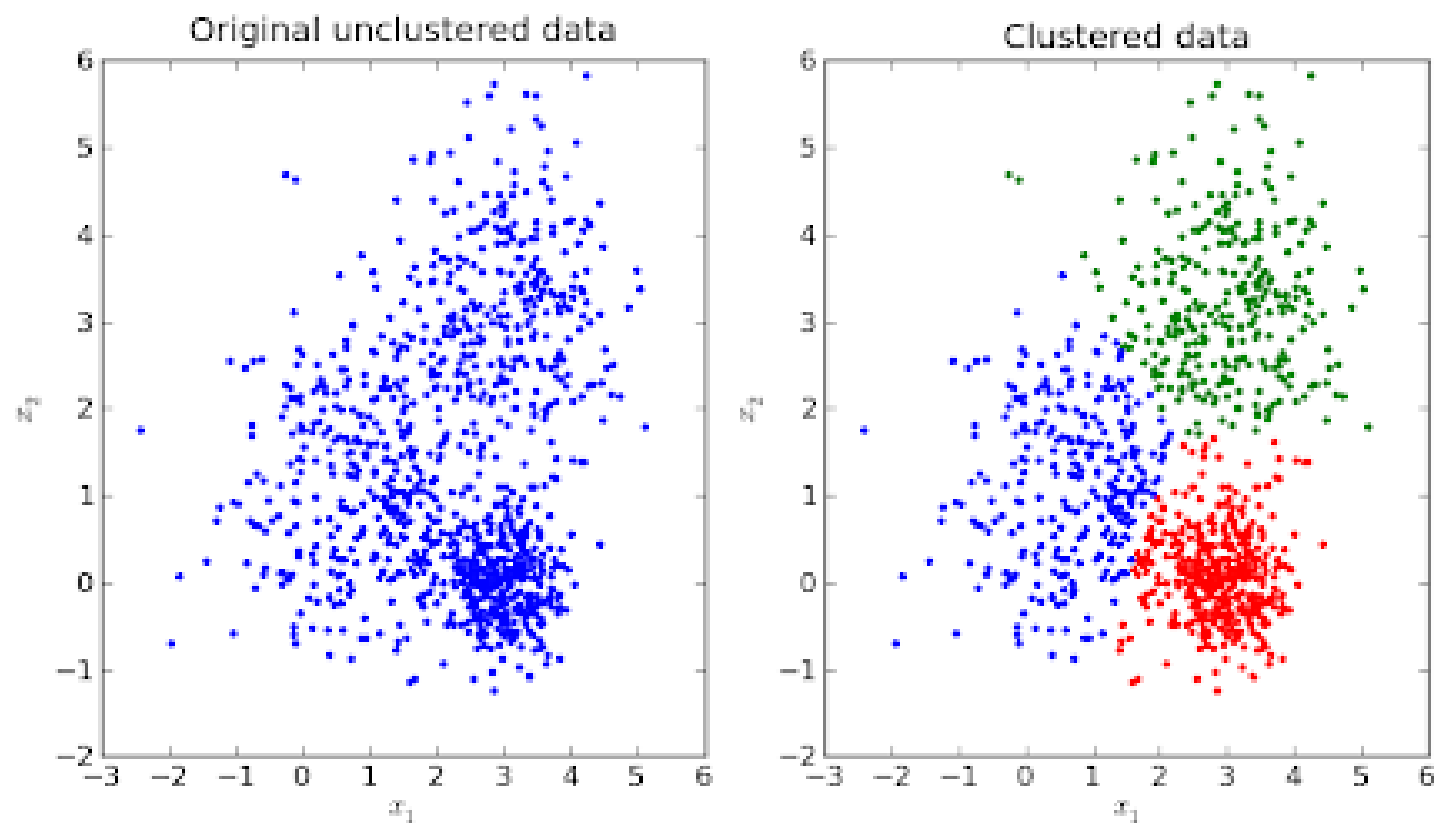
CLUSTERIZAÇÃO COM K-MEANS

K-MEANS

O k-means é um algoritmo de aprendizagem **não supervisionada**, ou seja, não necessita de dados rotulados para o seu aprendizado. O objetivo desse algoritmo é avaliar e agrupar dados de acordo com as suas características.

ALGUNS EXEMPLOS DE APLICAÇÕES COM K-MEANS:

- **Cientes:** produtos ou serviços semelhantes
- **Cientes:** características semelhantes
- **Séries/filmes:** gênero ou faixa etária
- **Usuários de uma rede social:** usuário influenciador
- **Paciente:** sintoma ou característica semelhante



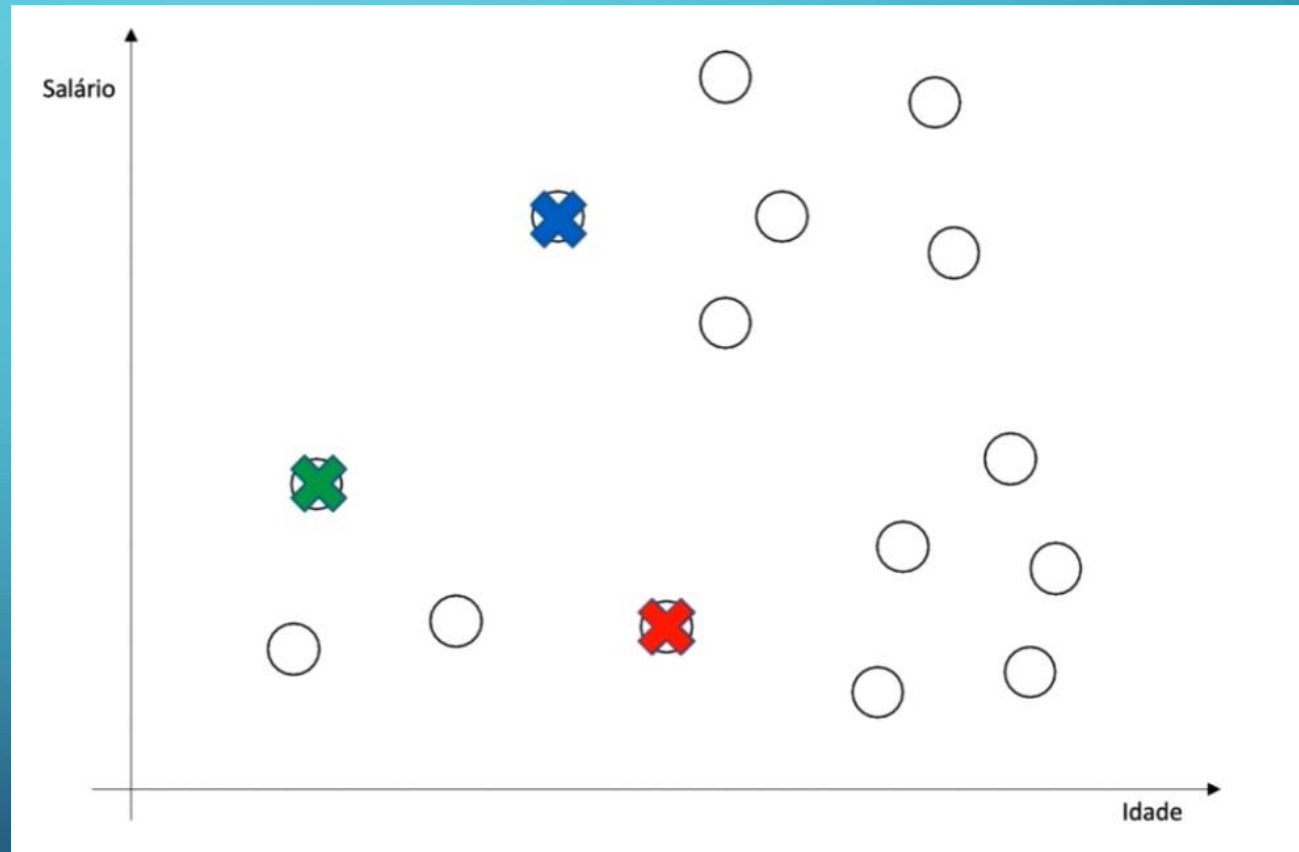
(Fonte: Clusterização de dados K-Means na biblioteca scikit-learn - Micreiros.com)

The background is a blue gradient with decorative white circuit-like lines in the corners. These lines consist of straight segments and small circles, resembling a stylized electronic circuit board.

EXEMPLO NO ORANGE

COMO O ALGORITMO TRABALHA:

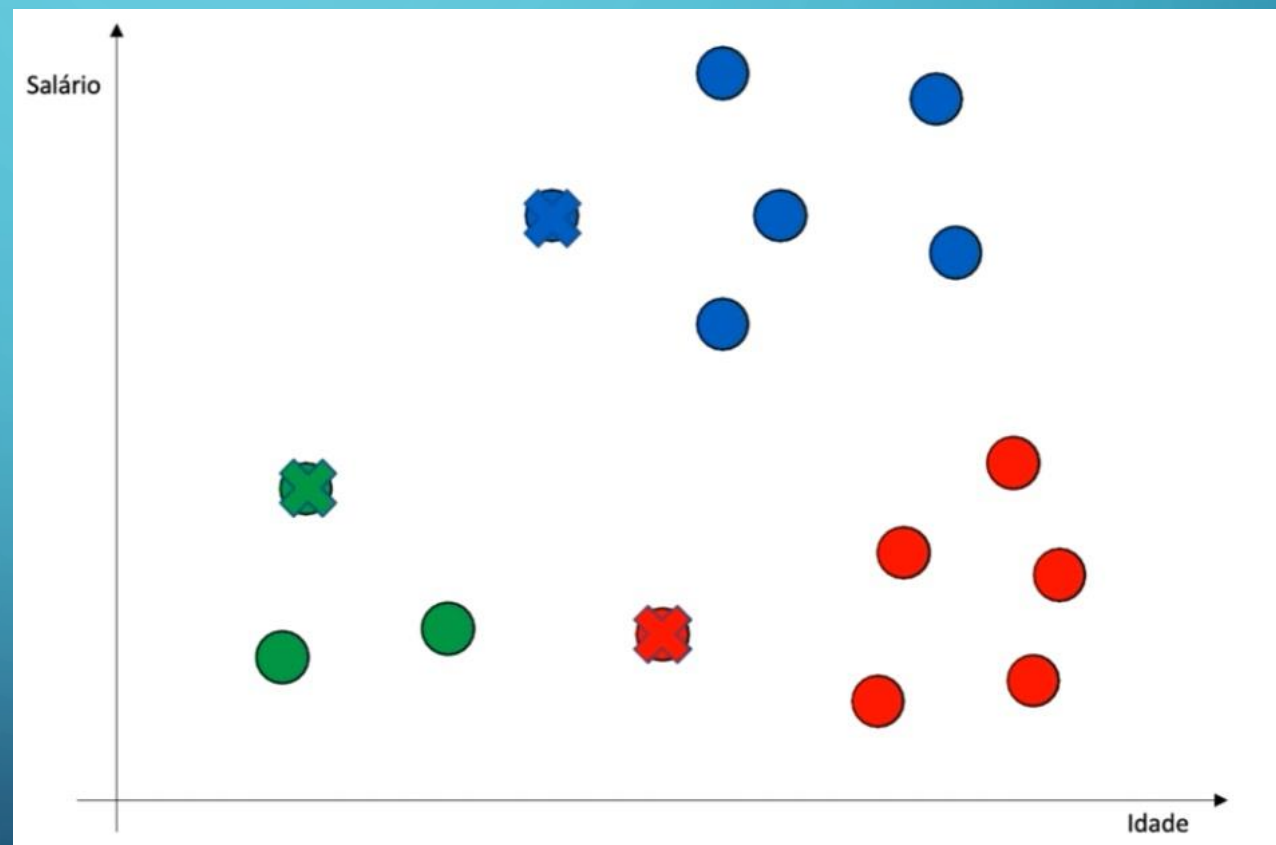
1. Primeiro, definimos um 'K', ou seja, um número de clusters (ou agrupamentos).
2. O algoritmo define, aleatoriamente, as posições dos centroides para cada cluster.
3. O próximo passo é calcular, para cada ponto, o centroide de menor distância. Cada ponto pertencerá ao centroide mais próximo.



(Machine Learning e Data Science: O Guia para Iniciantes | Udemy)

COMO O ALGORITMO TRABALHA:

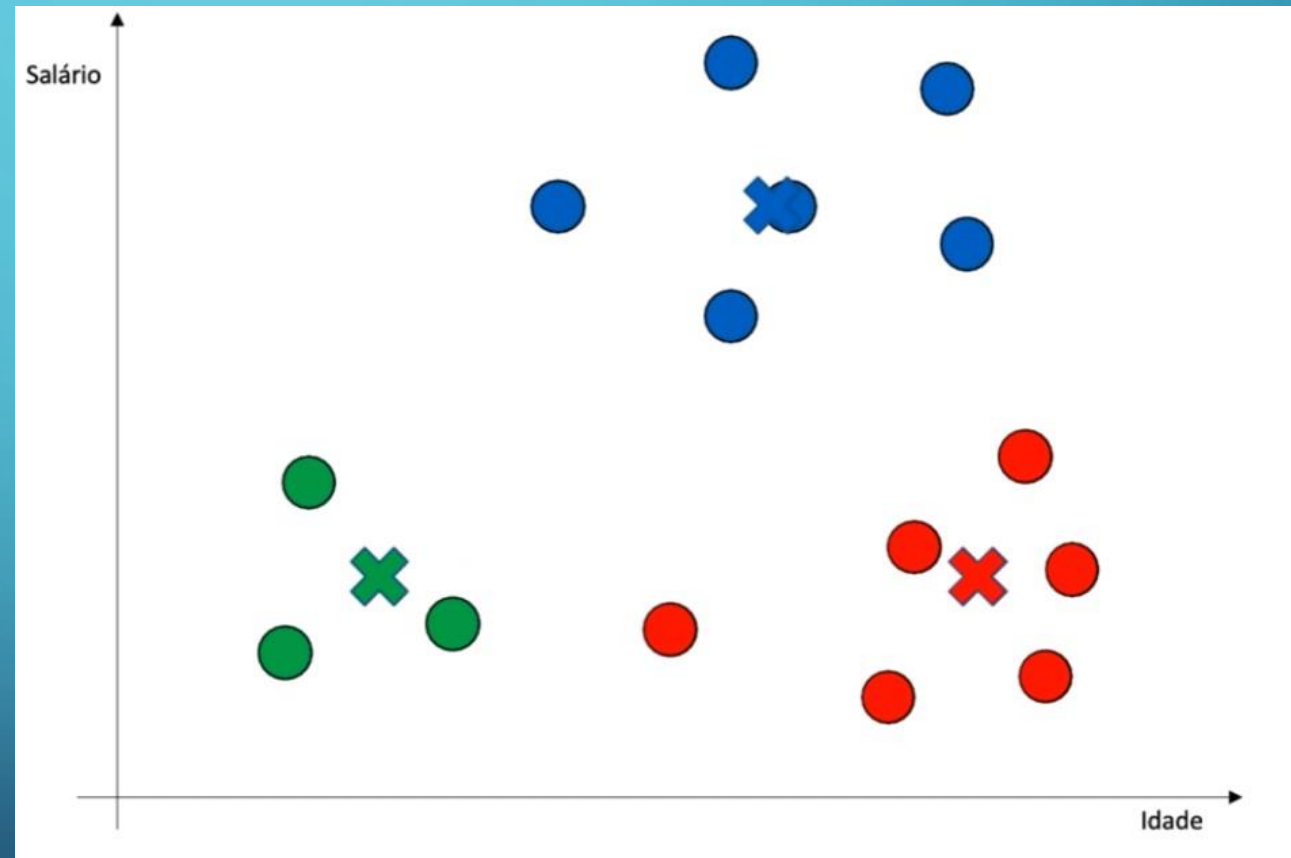
Os pontos serão classificados assim:



(Machine Learning e Data Science: O Guia para Iniciantes | Udemy)

COMO O ALGORITMO TRABALHA:

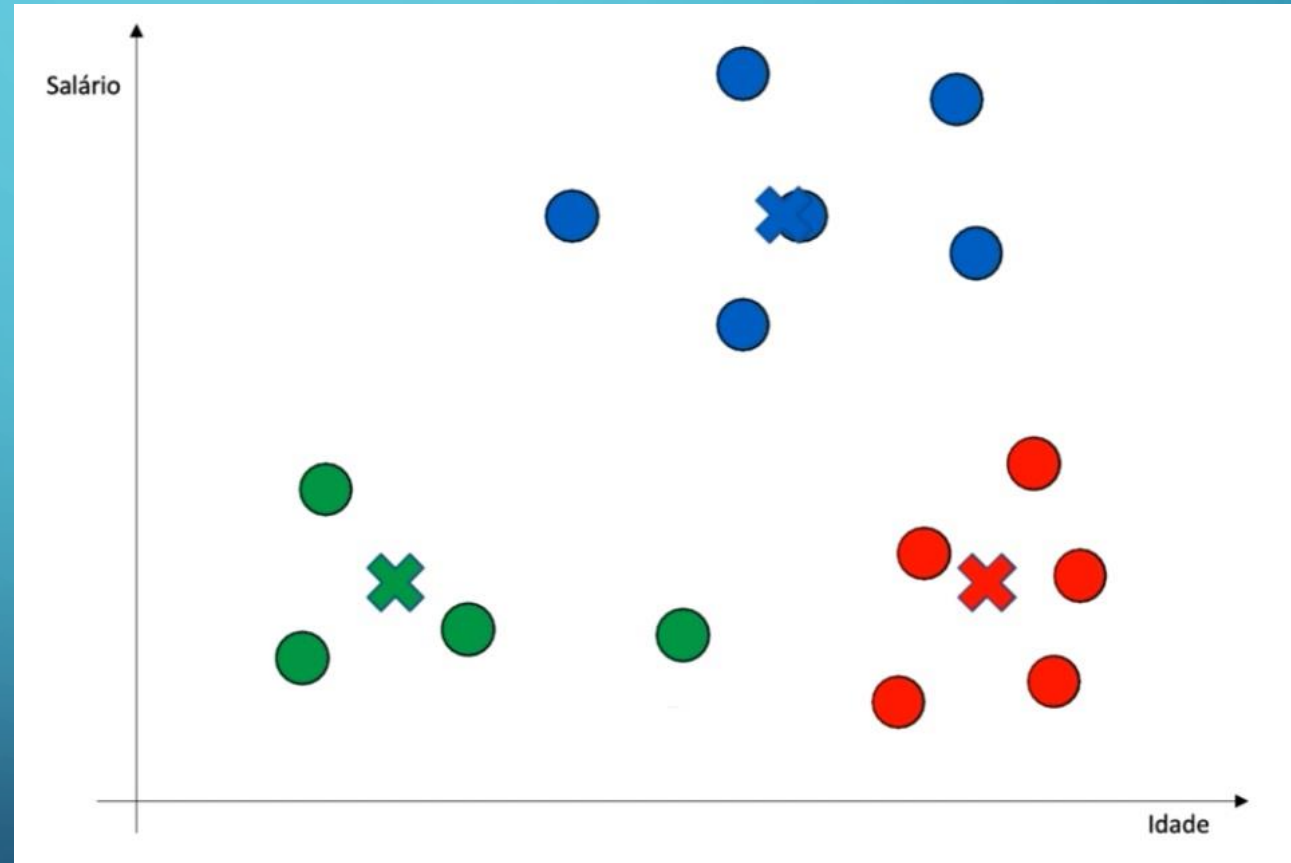
4. Agora, o algoritmo irá reposicionar o centroide. A nova posição do centroide deve ser a média da posição de todos os pontos do cluster.



(Machine Learning e Data Science: O Guia para Iniciantes | Udemy)

COMO O ALGORITMO TRABALHA:

5. Os dois últimos passos são repetidos, iterativamente, até o algoritmo encontrar a posição ideal dos centroides.



(Machine Learning e Data Science: O Guia para Iniciantes | Udemy)

REFERÊNCIAS BIBLIOGRÁFICAS + LINKS ÚTEIS

- [2.3. Clustering — scikit-learn 0.24.1 documentation](https://scikit-learn.org/stable/modules/clustering.html#k-means) (https://scikit-learn.org/stable/modules/clustering.html#k-means)
- [Clusterização de dados K-Means na biblioteca scikit-learn - Micreiros.com](https://micreiros.com/clusterizacao-de-dados-k-means-na-biblioteca-scikit-learn/) (https://micreiros.com/clusterizacao-de-dados-k-means-na-biblioteca-scikit-learn/)
- [sklearn.cluster.KMeans — scikit-learn 0.24.1 documentation](https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html?highlight=k%20means#sklearn.cluster.KMeans) (https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html?highlight=k%20means#sklearn.cluster.KMeans)
- [Machine Learning e Data Science: O Guia para Iniciantes | Udemy](https://www.udemy.com/course/guia-iniciantes-machine-learning-data-science/learn/lecture/21716962#overview) (https://www.udemy.com/course/guia-iniciantes-machine-learning-data-science/learn/lecture/21716962#overview)
- [K-means: o que é, como funciona, aplicações e exemplo em Python | by Bruno Anastacio | Programadores Ajudando Programadores | Medium](https://medium.com/programadores-ajudando-programadores/k-means-o-que-e-como-funciona-aplica%C3%A7%C3%B5es-e-exemplo-em-python-6021df6e2572) (https://medium.com/programadores-ajudando-programadores/k-means-o-que-e-como-funciona-aplica%C3%A7%C3%B5es-e-exemplo-em-python-6021df6e2572)