



Capacitação em Inteligência Artificial e Aplicações

Aprendizado Não Supervisionado:

- Prof. Gerson Vieira Albuquerque Neto
- Prof. Rodrigo Carvalho Souza Costa
- Prof. Yves Augusto Romero



UNIVERSIDADE
ESTADUAL DO CEARÁ



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO





IA

Planejamento da Disciplina

D	S	T	Q	Q	S	S
26	27 Introdução ao curso	28 Áreas e aplicações de IA	29 Tipos e definições de Inteligência artificial	30 Revisão de álgebra e probabilidade	31 Laboratório Python 1	1
2	3 Introdução aos classificadores supervisionados	4 Aula teórica Naive BaSim	5 Aula prática Naive BaSim	6 Feriado Semana Santa	7 Feriado Semana Santa	8
9	10 KNN + Métricas de Avaliação	11 Regressão Linear e Introdução à árvores de decisão	12 Prática Regressão Lienar + Árvores de Decisão	13 Introdução à Clusterização + KMédias	14 Introdução ao PCA / prática com classificadores já implementados	15
16	17 Introdução ao Perceptron Simples – Prática	18 Teoria MLP / Aplicação scilearn	19 Introdução ao DeepLearning	20 Uso de biblioteca DeepLearning	21 Feriado Tiradentes	28
23	24 Introdução ao TensorFlow / Keras	25 Introdução ao Pytorch	26 Tensorflow for android	27	28	29



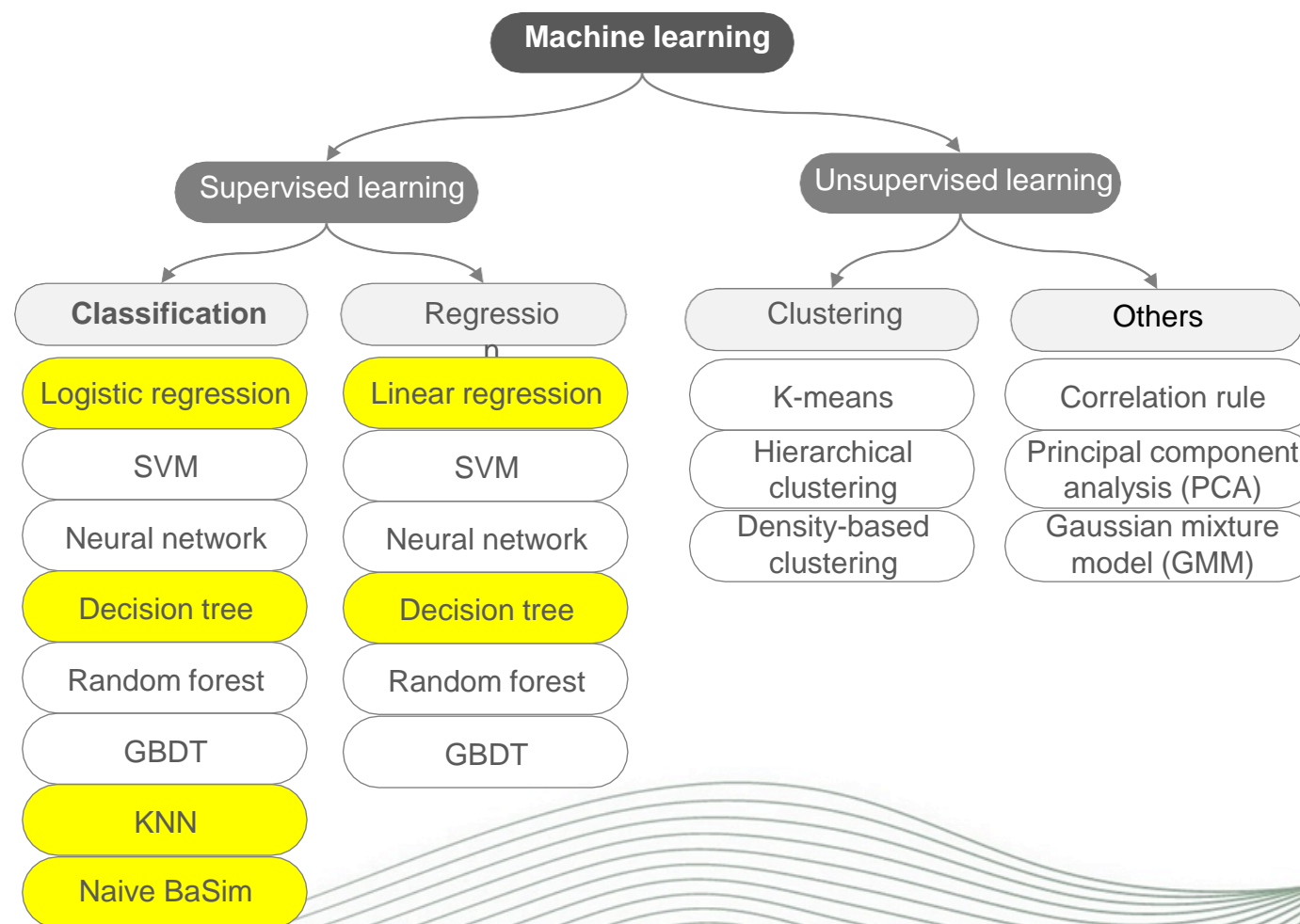
- Após a conclusão deste módulo, você será capaz de:
 - Compreender o funcionamento de algoritmos de clusterização
 - Compreender o funcionamento de métodos de redução de características
 - Compreender o conceito de Hiperparâmetros





IA

Revisão: Algoritmos de Machine Learning





Aprendizado Não Supervisionado:

- **Revisão sobre aprendizado não supervisionados**
- Clustering
- Redução de dimensionalidade
- Hiperparâmetros



UNIVERSIDADE
ESTADUAL DO CEARÁ



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO



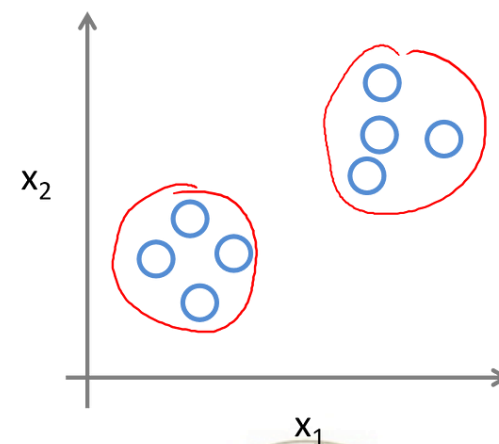


IA

Revisão clusterização

- O agrupamento (**clustering**) é uma forma comum de aprendizagem não supervisionada.
- Para amostras não rotuladas, os algoritmos de aprendizado modelam diretamente os conjuntos de dados de entrada.
- Precisamos apenas colocar amostras altamente semelhantes juntas, calcular a semelhança entre amostras novas e as existentes e classificá-las por semelhança.

Unsupervised Learning

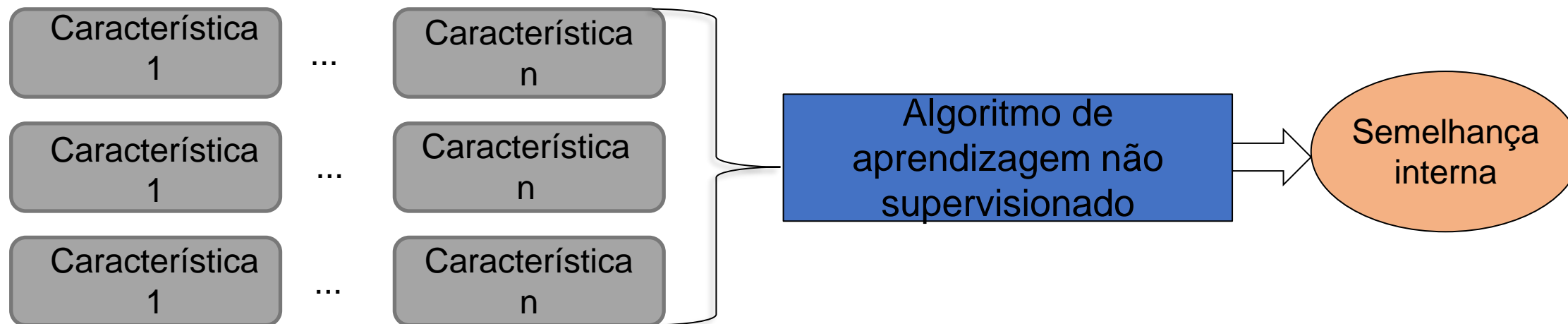




IA

Aprendizagem não-supervisionada

Espaço de Dados



Consumo Mensal	Mercadoria	Tempo de Consumo
1000–2000	Raquete de badminton	6:00–12:00
500–1000	Basquete	18:00–24:00
1000–2000	Console de jogos	00:00–6:00

Categoria
Cluster 1
Cluster 2



IA

Revisão





Aprendizado Não Supervisionado:

- Revisão de aprendizado não supervisionado
- **Clustering**
- Redução de dimensionalidade
- Hiperparâmetros



UNIVERSIDADE
ESTADUAL DO CEARÁ

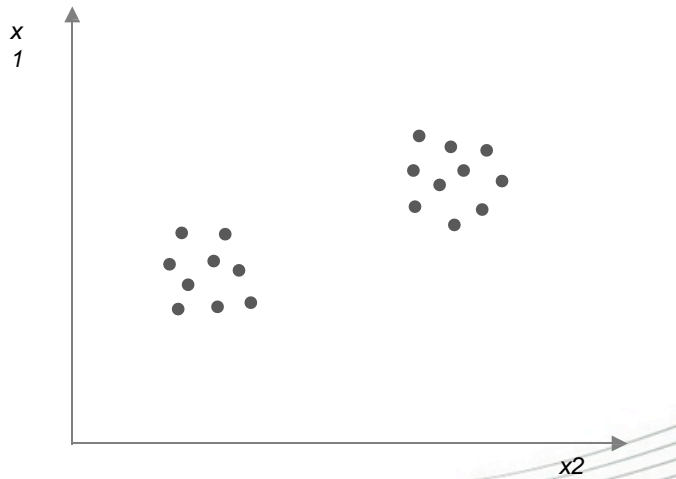


MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO

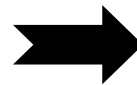




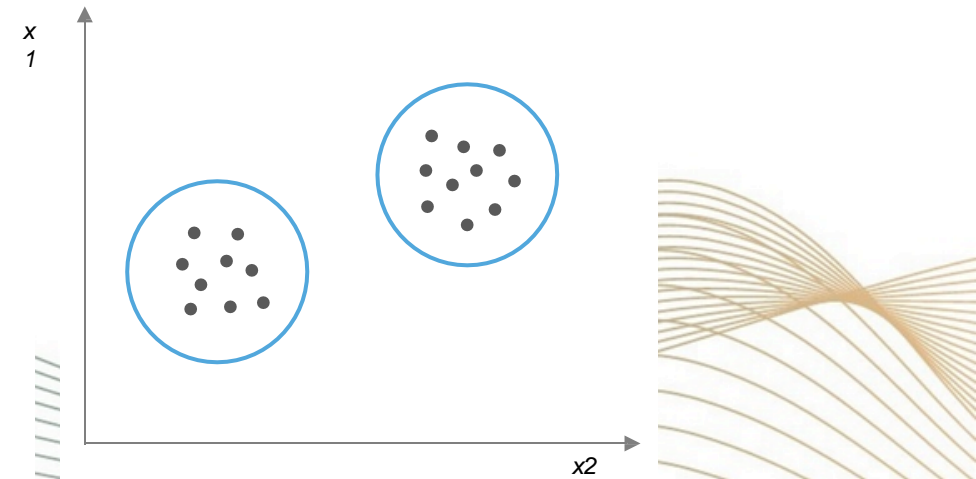
- O agrupamento K-médias visa particionar n observações em k clusters em que cada observação pertence ao cluster com a média mais próxima, servindo como um protótipo do cluster.
- Para o algoritmo K-médias, especifique o número final de clusters (k). Em seguida, divida n objetos de dados em k clusters.
- Os clusters obtidos atendem às seguintes condições:
 1. Os objetos no mesmo cluster são altamente semelhantes.
 2. A semelhança de objetos em diferentes aglomerados é pequena.



K-means clustering



Os dados **não está rotulado**. O cluster K-médias pode classificar automaticamente os conjuntos de dados.

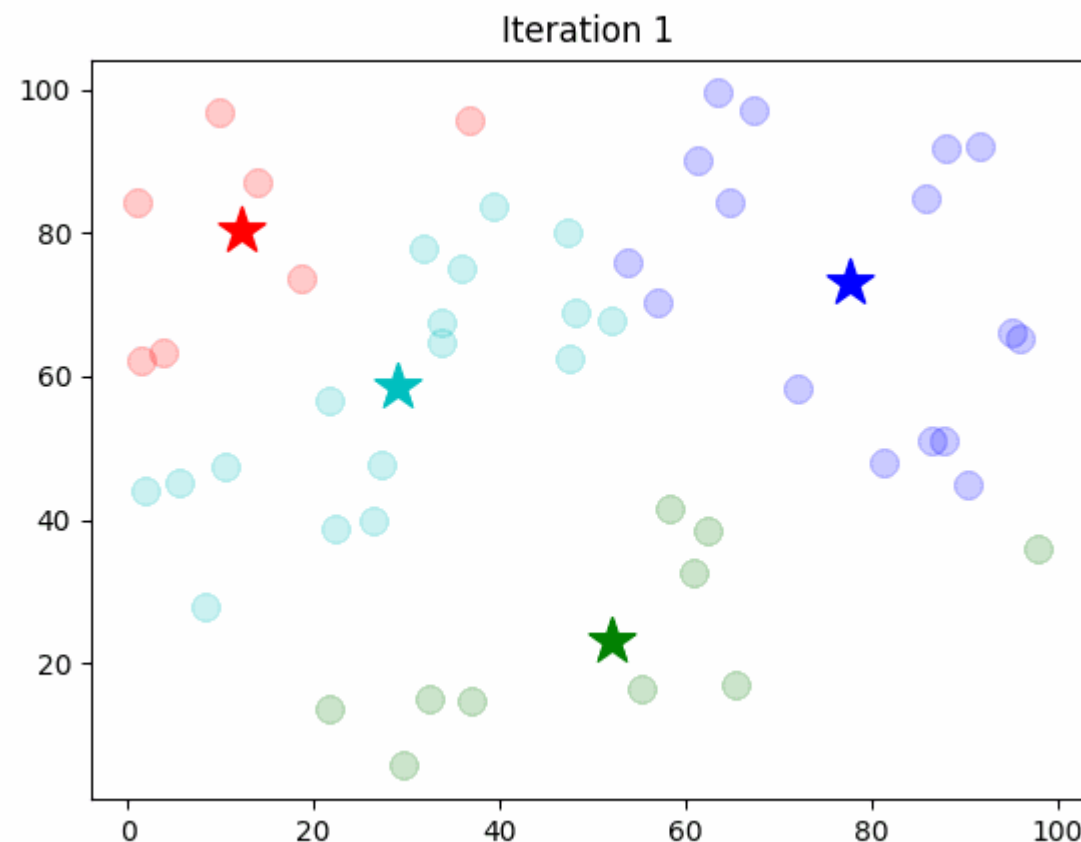




IA

Procedimento Básico K-médias

1. Especifica-se o número K de clusters
2. Inicializa-se aleatoriamente os centroids.
3. Repita até que a posição dos centroids não mude ou o número máximo de iterações seja alcançado
 1. Realiza-se a classificação do dataset em função do centroid mais próximo
 2. Recalcula-se o valor do centroid dos de cada cluster como a média dos pontos do cluster



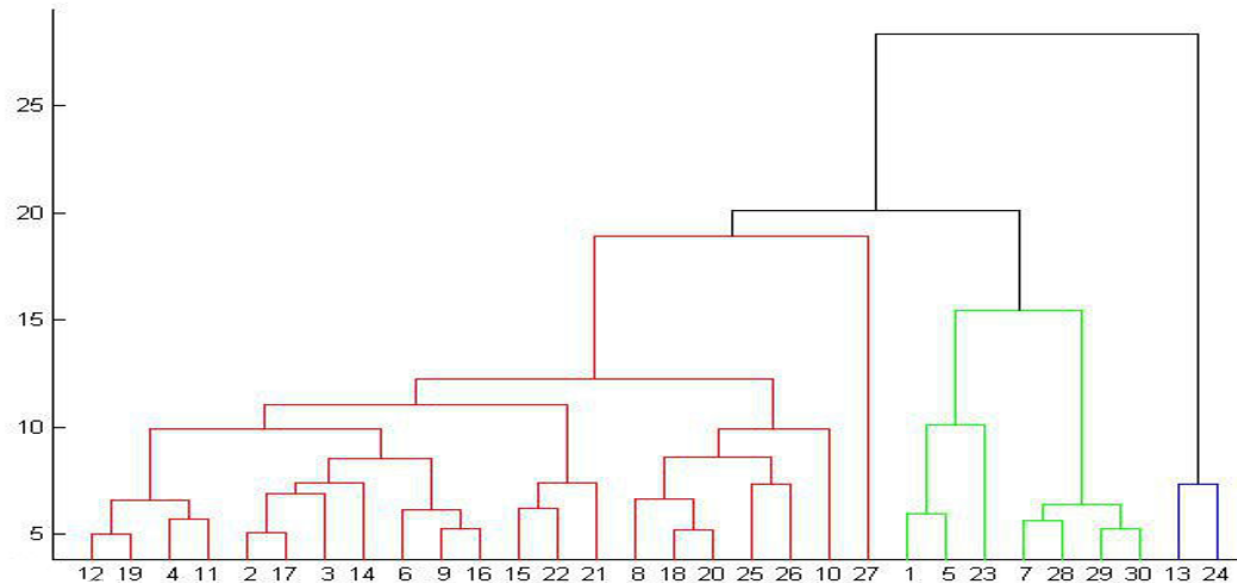


IA

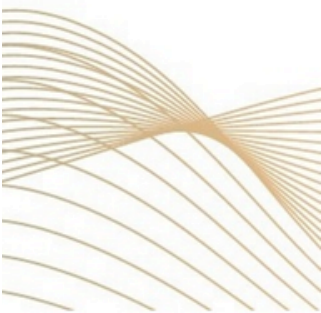
Agrupamento Hierárquico

- O clustering hierárquico divide um conjunto de dados em diferentes camadas e forma uma estrutura de clustering semelhante a uma árvore.
- A divisão do conjunto de dados pode usar uma política de agregação "de baixo para cima" ou uma política de divisão de "cima para baixo". A hierarquia de agrupamento é representada em um gráfico de árvore.
- A raiz é o aglomerado único de todas as amostras, e as folhas são o aglomerado de apenas uma amostra.

Agglomerative
hierarchical clustering



Split hierarchical
clustering





Aprendizado Não Supervisionado:

- Revisão de aprendizado não supervisionado
- Clustering
- **Redução de dimensionalidade**
- Hiperparâmetros



UNIVERSIDADE
ESTADUAL DO CEARÁ



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO

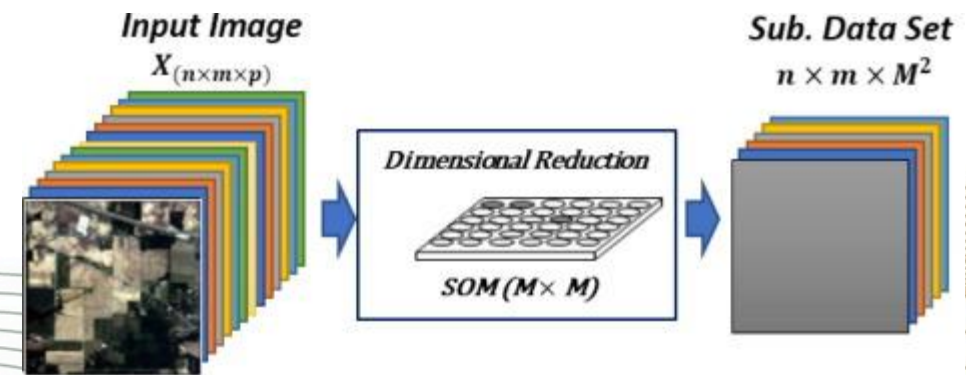
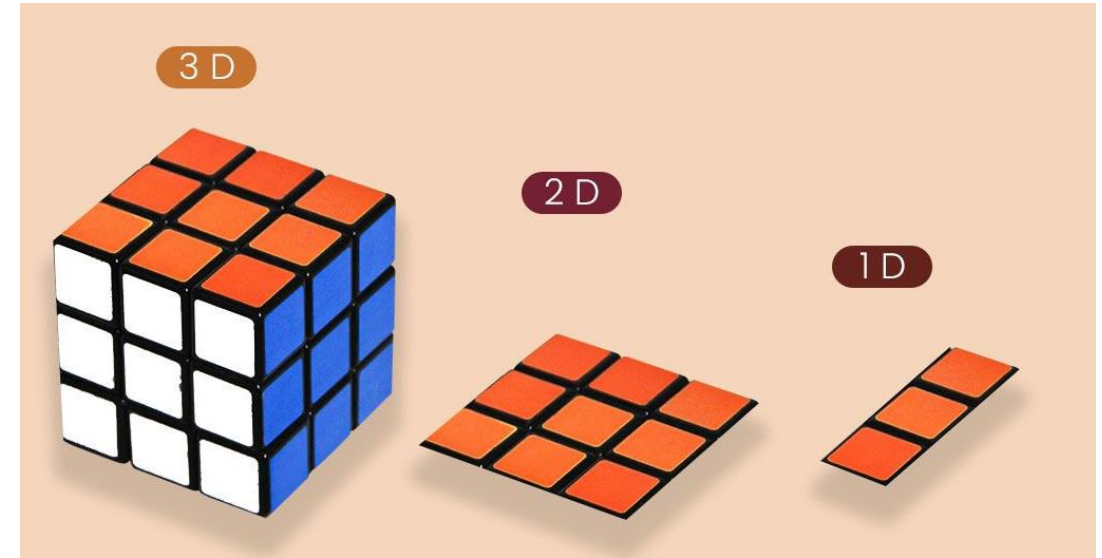




IA

Redução de Dimensionalidade

- Transformação de dados de um espaço de alta dimensão em um espaço de baixa dimensão
- Idealmente perto de sua dimensão intrínseca.
- Métodos
 - Seleção de atributos
 - Projeção de atributos

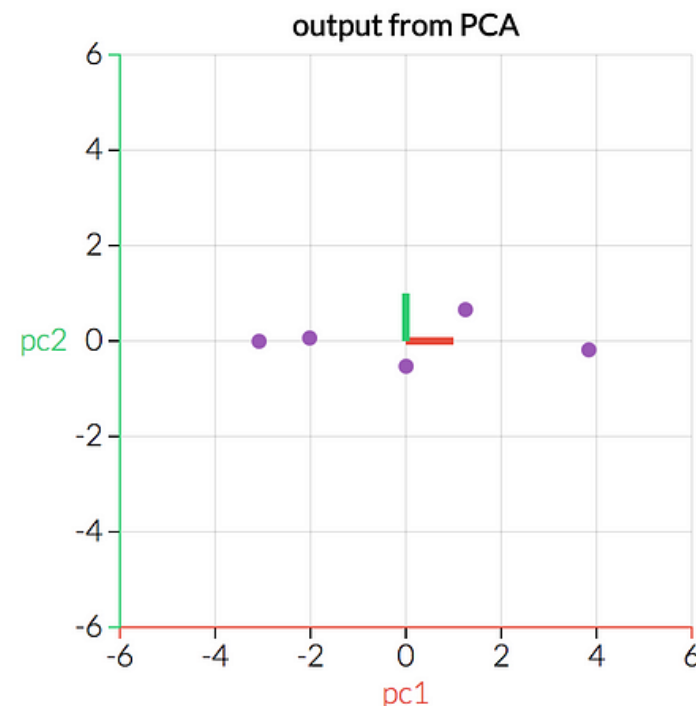
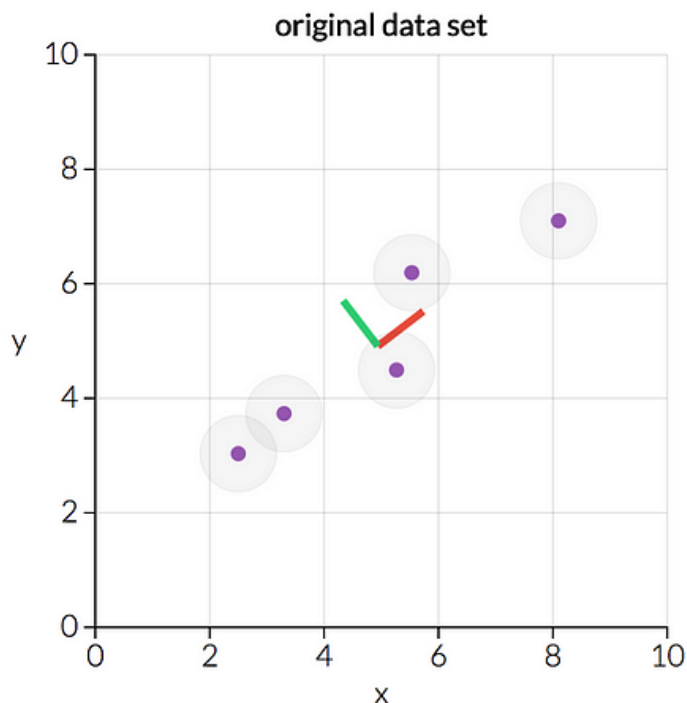




IA

Análise de Componentes Principais PCA - Principal Component Analysis

- O PCA é frequentemente usado para reduzir a dimensionalidade de grandes conjuntos de dados, transformando um grande conjunto de variáveis em um menor que ainda contém a maioria das informações no grande conjunto.

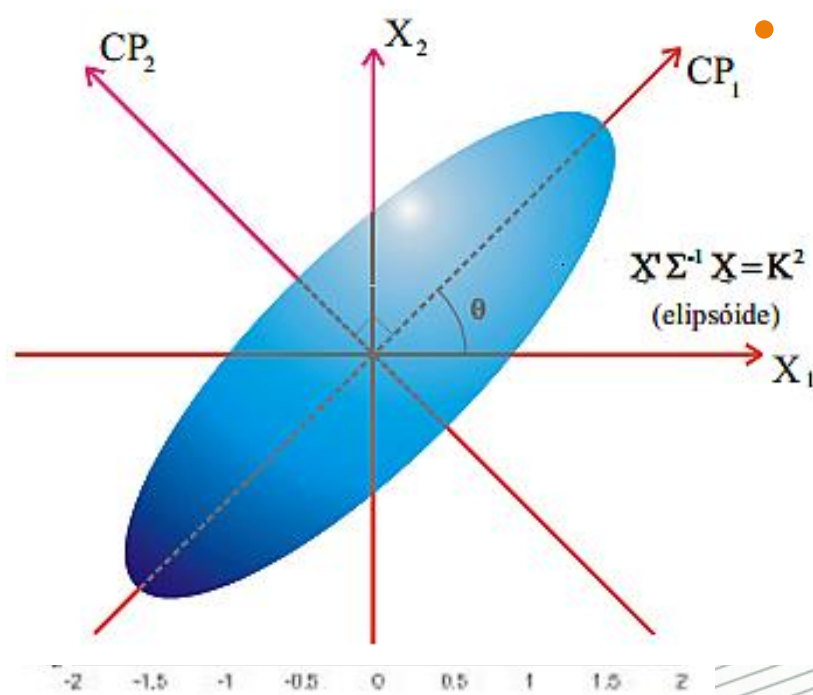




IA

Componentes Principais

- Consiste em uma técnica estatística que realiza uma transformação linear de um conjunto de variáveis originais em um menor conjunto com variáveis descorrelacionadas, representando a maior parte da informação do conjunto de dados originais.



- A Componente Principal é o arranjo que melhor representa a distribuição dos dados (**linha vermelha**) e a Componente secundária é perpendicular a componente principal (**linha azul**).



- O c-componente principal é encontrado a partir da transformação do espaço original através do autovetor $E_c = (e_1, e_2, \dots, e_n)$ associado ao c-maior autovalor da matriz de covariância C_X , ou seja:

$$y = x \cdot E_c.$$

X1 →
X2 →
X3 →
⋮
Xp →

P - variáv

Calcular matriz
de covariância
 C_X



Encontrar os
autovalores
 Λ



Encontrar os c
auto-vetores
 E_c



Seleção de
novas
características

→ Y1
→ Y2
→ Y3
⋮
→ Yp

*P - componentes
Principais*



- A matriz de covariância (C_X) do espaço característico X é definida como:

$$C_X = \sum_{i=1}^m (x_i - \bar{x}) (x_i - \bar{x})^T$$

$$C = \begin{pmatrix} cov(x, x) & cov(x, y) & cov(x, z) \\ cov(y, x) & cov(y, y) & cov(y, z) \\ cov(z, x) & cov(z, y) & cov(z, z) \end{pmatrix}$$

- Em que:
 - Os componentes c_{ij} da matriz de covariância (C_X) representam a covariância entre as variáveis i e j .
 - Quando duas características j e k dos dados estão correlacionadas, sua covariância é nula ($c_{jk} = c_{kj} = 0$).
- Escolhendo os autovetores associados aos maiores autovalores, há a menor perda possível de informação neste novo espaço característico



IA

Calculando os autovalores e autovetores de C_x

- Para calcular os autovalores λ da matriz C_x deve-se resolver a equação característica:

$$\det(C_x - \lambda I) = 0$$

- Sabendo que para cada autovalor λ_i encontrado, resolvemos o sistema linear $(C_x - \lambda I)v = 0$ para calcular o autovetor v associado ao autovalor λ_i

x	y
2	3
2	4
3	4
4	4
5	4
5	5
6	6
6	7
7	7
8	7

\bar{x}	\bar{y}
4,8	5,1

\bar{x}

37,6	25,2
25,2	20,9

C_x

$37,6 - \lambda$	25,2
25,2	$20,9 - \lambda$

$C_x - \lambda I$

$$\lambda^2 - 58,5\lambda + 150,8 = 0$$

$$\det(C_x - \lambda I) = 0$$



IA

Calculando os autovalores e autovetores de C_x

- Para calcular os autovalores λ da matriz C_x deve-se resolver a equação característica:

$$\det(C_x - \lambda I) = 0$$

- Sabendo que para cada autovalor λ_i encontrado, resolvemos o sistema linear $(C_x - \lambda I)v = 0$ para calcular o autovetor v associado ao autovalor λ_i

x	y
2	3
2	4
3	4
4	4
5	4
5	5
6	6
6	7
7	7
8	7

55,7974	0
0	2,70264

autovalores

$$E_1 = 0.81071949x - 0.58543481y$$

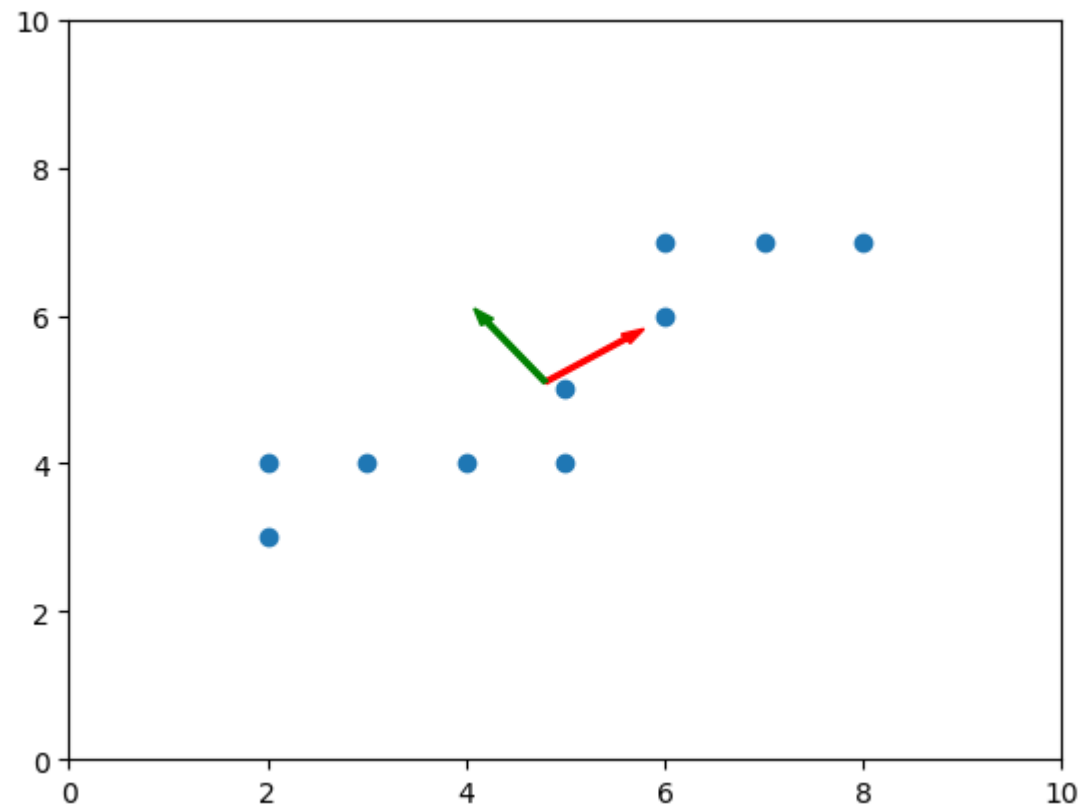
$$E_2 = 0.58543481x + 0.81071949y$$



IA

Exemplo

$$E_1 = 0.81071949x - 0.58543481y$$
$$E_2 = 0.58543481x + 0.81071949y$$





Aprendizado Não Supervisionado:

- Revisão de aprendizado não supervisionado
- Clustering
- Redução de dimensionalidade
- **Hiperparâmetros**



UNIVERSIDADE
ESTADUAL DO CEARÁ



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO

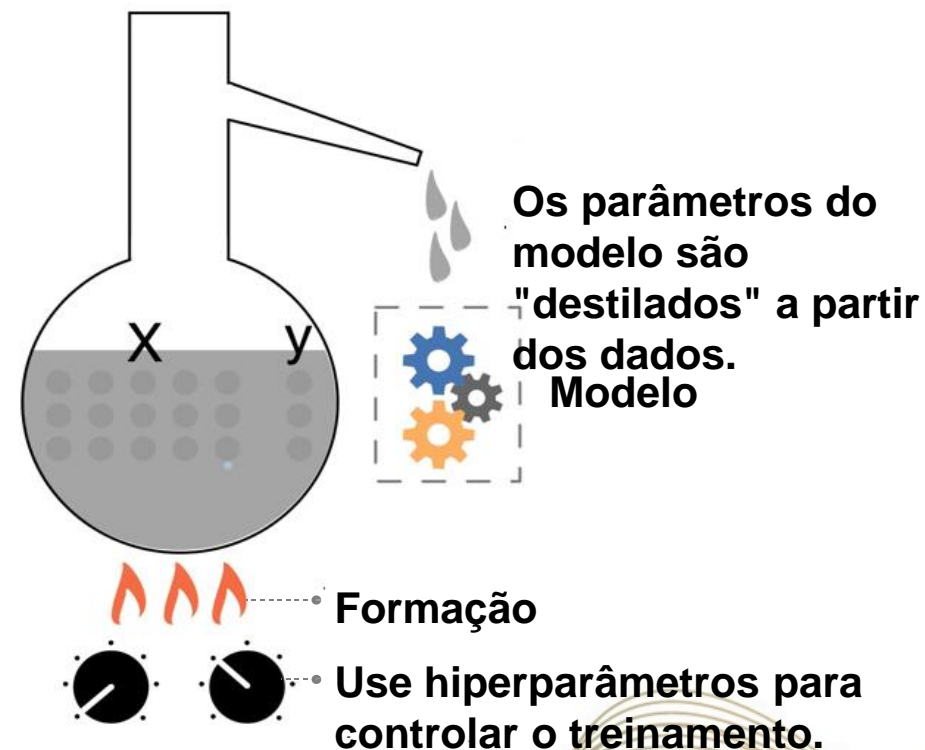




IA

Parâmetros e hiperparâmetros em modelos

- O modelo contém não apenas parâmetros, mas também hiperparâmetros.
- O objetivo é permitir que o modelo aprenda os parâmetros ideais.
 - Os parâmetros são aprendidos automaticamente por modelos.
 - Os hiperparâmetros são definidos manualmente.





IA

Hiperparâmetros de um modelo

- Frequentemente usado em processos de estimativa de parâmetros de modelo.
- Muitas vezes especificado pelo praticante.
- Muitas vezes pode ser definido usando heurísticas.
- Muitas vezes ajustado para um determinado problema de modelagem preditiva.

Os hiperparâmetros do modelo são configurações externas dos modelos.

- λ durante a regressão de Lasso/Ridge
- Taxa de aprendizado para treinar uma rede neural, número de iterações, tamanho do lote, função de ativação e número de neurônios
- C e σ em máquinas vetoriais de suporte (SVM)
- K em k-vizinho mais próximo (KNN)
- Número de árvores em uma floresta aleatória

Hiperparâmetros comuns do modelo



Procedimento para
pesquisar
hiperparâmetros

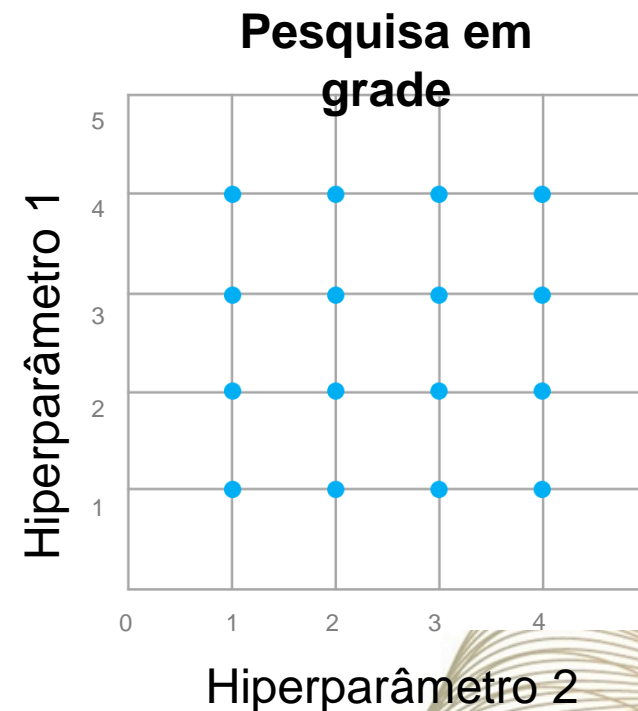
1. Dividir um conjunto de dados em um conjunto de treinamento, conjunto de validação e conjunto de testes.
2. Otimização dos parâmetros do modelo usando o conjunto de treinamento com base nos indicadores de desempenho do modelo.
3. Procurar os hiperparâmetros do modelo usando o conjunto de validação com base nos indicadores de desempenho do modelo.
4. Execute as etapas 2 e 3 alternadamente. Finalmente, determine os parâmetros e hiperparâmetros do modelo e avalie o modelo usando o conjunto de testes.

Algoritmo de busca
(etapa 3)

- Busca em grade
- Busca aleatória
- Busca inteligente heurística
- Busca bayesiana



- A pesquisa em grade tenta pesquisar exaustivamente todas as combinações de hiperparâmetros possíveis para formar uma grade de valores de hiperparâmetros.
- Na prática, o intervalo de valores de hiperparâmetros a serem pesquisados é especificado manualmente.
- A pesquisa em grade é um método caro e demorado.
- Este método funciona bem quando o número de hiperparâmetros é relativamente pequeno. Portanto, é aplicável a algoritmos de aprendizado de máquina em geral, mas inaplicável a redes neurais (consulte a parte de aprendizado profundo).



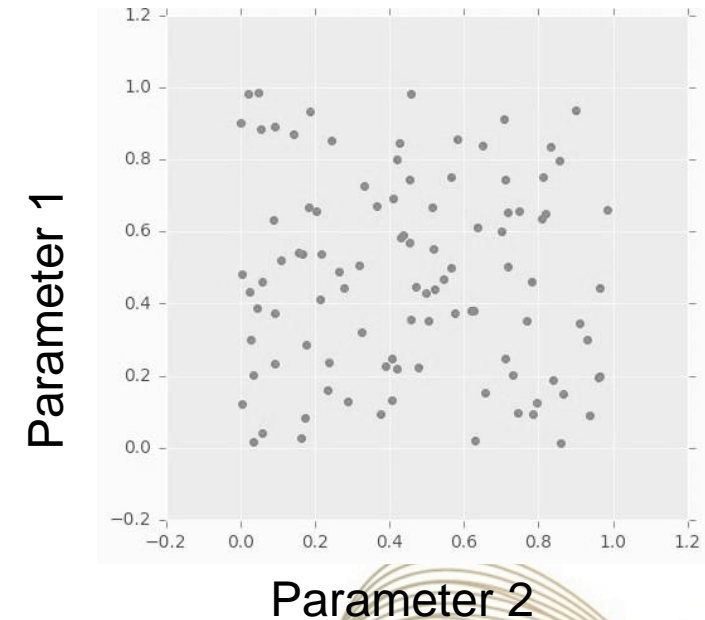


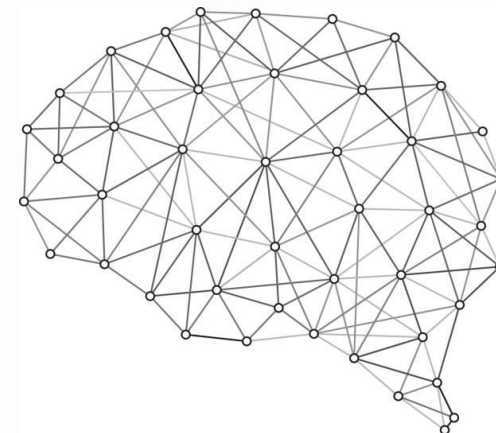
IA

Método de pesquisa de hiperparâmetros - Pesquisa aleatória

- Quando o espaço de pesquisa de hiperparâmetro é grande, a pesquisa aleatória é melhor do que a pesquisa em grade.
- Na pesquisa aleatória, cada configuração é amostrada a partir da distribuição de possíveis valores de parâmetros, na tentativa de encontrar o melhor subconjunto de hiperparâmetros.
- Nota:
 - A pesquisa é realizada dentro de um intervalo grosseiro, que será reduzido com base em onde o melhor resultado aparece.
 - Alguns hiperparâmetros são mais importantes do que outros, e o desvio de pesquisa será afetado durante a pesquisa aleatória.

Random search





Classificadores Supervisionados: próxima aula

- Prática de aprendizado não supervisionado



UNIVERSIDADE
ESTADUAL DO CEARÁ



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO



Dúvidas?

Módulo de Inteligência Artificial



UNIVERSIDADE
ESTADUAL DO CEARÁ



Instituto Iracema
PESQUISA E INOVAÇÃO



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO

