

**Educação  
Profissional  
Paulista**

Técnico em  
**Ciência de  
Dados**

# Lógica de programação e algoritmos

## Decomposição de problemas

Aula 1

Código da aula: [DADOS]ANO1C3B3S19A1

Lógica de  
programação e  
algoritmos

# Mapa da unidade 1 Componente 4

semana  
**21**

Recursividade

semana  
**18**

Prática de  
decomposição de  
problemas

semana  
**19**

**Você está aqui!**

Decomposição de problemas

semana  
**23**

Busca e  
ordenação

Lógica de  
programação e  
algoritmos

## Mapa da unidade 1 Componente 4

**Você está aqui!**

Decomposição de problemas

19

**Aula 1**

Código da aula: [DADOS]ANO1C3B3S19A1



## Objetivo da aula

- Conhecer a importância de quebrar problemas complexos em partes menores e mais gerenciáveis, facilitando, assim, a análise e a solução.



## Recursos didáticos

- Recurso audiovisual para exibição de vídeos e imagens;
- Acesso ao laboratório de informática e/ou internet.



## Duração da aula

50 minutos.



## Competências técnicas

- Usar técnicas para explorar e analisar dados, aplicar modelos estatísticos, identificar padrões, realizar inferências e tomar decisões baseadas em evidências;
- Compreender e dominar técnicas de manipulação de dados; extrair, transformar e carregar conjuntos de dados de diferentes fontes, garantindo a qualidade e a integridade dos dados; criar e compreender visualizações gráficas.



## Competências socioemocionais

- Colaborar efetivamente com outros profissionais, como cientistas de dados e engenheiros de dados;
- trabalhar em equipes multifuncionais colaborando com colegas, gestores e clientes.







## Primeiras ideias

Como você pode simplificar a solução de um problema muito complicado dividindo-o em partes menores?



Quais são os benefícios de resolver pequenas partes de um problema antes de tentar resolver tudo de uma vez?



Como você pode identificar as diferentes partes de um problema grande?



Como determinar quais partes de um problema são mais urgentes ou críticas para resolver primeiro, ao dividir um grande desafio em componentes menores?

## Ponto de partida

Como decompor o problema de aumentar o número de clientes do *E-Commerce Boost*?

Imagine que você é um técnico em ciência de dados em uma empresa de comércio eletrônico e sua equipe está trabalhando para aumentar as vendas. Você recebeu um conjunto de dados com informações sobre os clientes, como idade, gênero, histórico de compras e valor gasto.

- ▶ Seu objetivo é criar um modelo que preveja quais clientes têm maior probabilidade de realizar compras novamente.
- ▶ Como você resolveria esse problema?

Na próxima seção e nas próximas aulas, mergulharemos no problema, tentaremos definir juntos a melhor estratégia de decomposição de problemas para este estudo de caso.

Situação fictícia elaborada especialmente para o curso.



Construindo  
o **conceito**

## Como uma empresa pode deixar bem claro quais são seus problemas para encontrar as melhores soluções?

A Ciência de Dados é essencial nos dias de hoje, pois permite às empresas tomarem decisões mais informadas e assertivas. Ela ajuda a:

- ▶ entender padrões;
- ▶ prever tendências;
- ▶ identificar oportunidades de negócio.

Problemas mal definidos podem levar a análises equivocadas, modelos de baixa qualidade e, conseqüentemente, a decisões erradas.

Isso pode resultar em atrasos na evolução da empresa e, até mesmo, em prejuízos financeiros significativos.

Construindo  
o **conceito**

# Técnico em Ciência de Dados

O que se espera de um técnico em Ciência de Dados em comparação a um cientista de dados?

## **Técnico em Ciência de Dados**

É esperado que ele possua um entendimento básico dos conceitos abordados nas aulas.

## **Cientista de dados com formação superior**

É esperado que tenha um conhecimento mais aprofundado e uma capacidade analítica avançada para aplicar esses conceitos na resolução de problemas complexos e na extração de insights, a partir de grandes volumes de dados.

Construindo  
o **conceito**

# Introdução

Para entendermos como funciona a decomposição de problemas em Ciência de Dados, será necessário abordar os tipos de problemas existentes que a Ciência de Dados busca resolver. Vamos lá!



Reprodução - GUESSE, 2023.



## Construindo o conceito

# Problema de classificação

O objetivo é prever a qual categoria ou classe uma nova observação pertence, com base em observações anteriores rotuladas.

Exemplos:

- prever se um e-mail é spam ou não;
- Imagine que você tem várias frutas, como maçãs, bananas e laranjas, e quer que um computador aprenda a identificar qual fruta é qual. Isso é um problema de classificação. O computador vai aprender a separar as frutas em diferentes grupos com base em características como cor, formato e tamanho.



Reprodução – GUESSE, 2023.

## Construindo o **conceito**

# Problema de regressão

O objetivo é prever um valor contínuo para uma nova observação, com base em observações anteriores.

Exemplos:

- prever o preço de uma casa com base em suas características;
- prever o preço de uma casa com base em características como o número de quartos, o tamanho do terreno e a localização. Esse é um problema de regressão. O computador vai aprender a encontrar uma relação entre essas características e o preço da casa para fazer previsões precisas.



Reprodução – GUESSE, 2023.

## Problema de agrupamento

O objetivo é agrupar um conjunto de observações em grupos ou *clusters*, em que observações dentro de um cluster são mais semelhantes entre si do que com observações de outros *clusters*.

Exemplo:

- agrupar clientes com base em seus padrões de compra. Imagine que você tem um monte de dados sobre clientes de uma loja e quer agrupá-los em diferentes grupos, com base em seus hábitos de compra. Isso é um problema de agrupamento. O computador vai aprender a identificar padrões nos dados para agrupar os clientes de forma inteligente.



Reprodução – GUESSE, 2023.





Vamos  
fazer um  
**quiz**

## Qual é o objetivo do agrupamento (*clustering*) em Aprendizagem de Máquina?

Identificar relações frequentes  
entre variáveis.

Agrupar observações em  
grupos semelhantes.

Prever a categoria ou classe  
de uma nova observação.

Reduzir a quantidade de  
variáveis em um conjunto de  
dados.



Vamos  
fazer um  
**quiz**

## Qual é a principal diferença entre aprendizado supervisionado e aprendizado não supervisionado?

No aprendizado supervisionado os dados  
são rotulados; no não supervisionado os  
dados não são rotulados.

No aprendizado supervisionado não  
há modelo de aprendizado; no não  
supervisionado há um modelo de  
aprendizado.

No aprendizado supervisionado os  
dados não são rotulados; no não  
supervisionado os dados são rotulados.

No aprendizado supervisionado não há  
necessidade de conjunto de dados; no  
não supervisionado o conjunto de dados  
é essencial.





Vamos  
fazer um  
**quiz**

## Qual é a principal diferença entre o problema de regressão e o problema de classificação em Aprendizagem de Máquina?

O problema de regressão visa prever valores contínuos, enquanto o problema de classificação visa prever classes ou categorias.

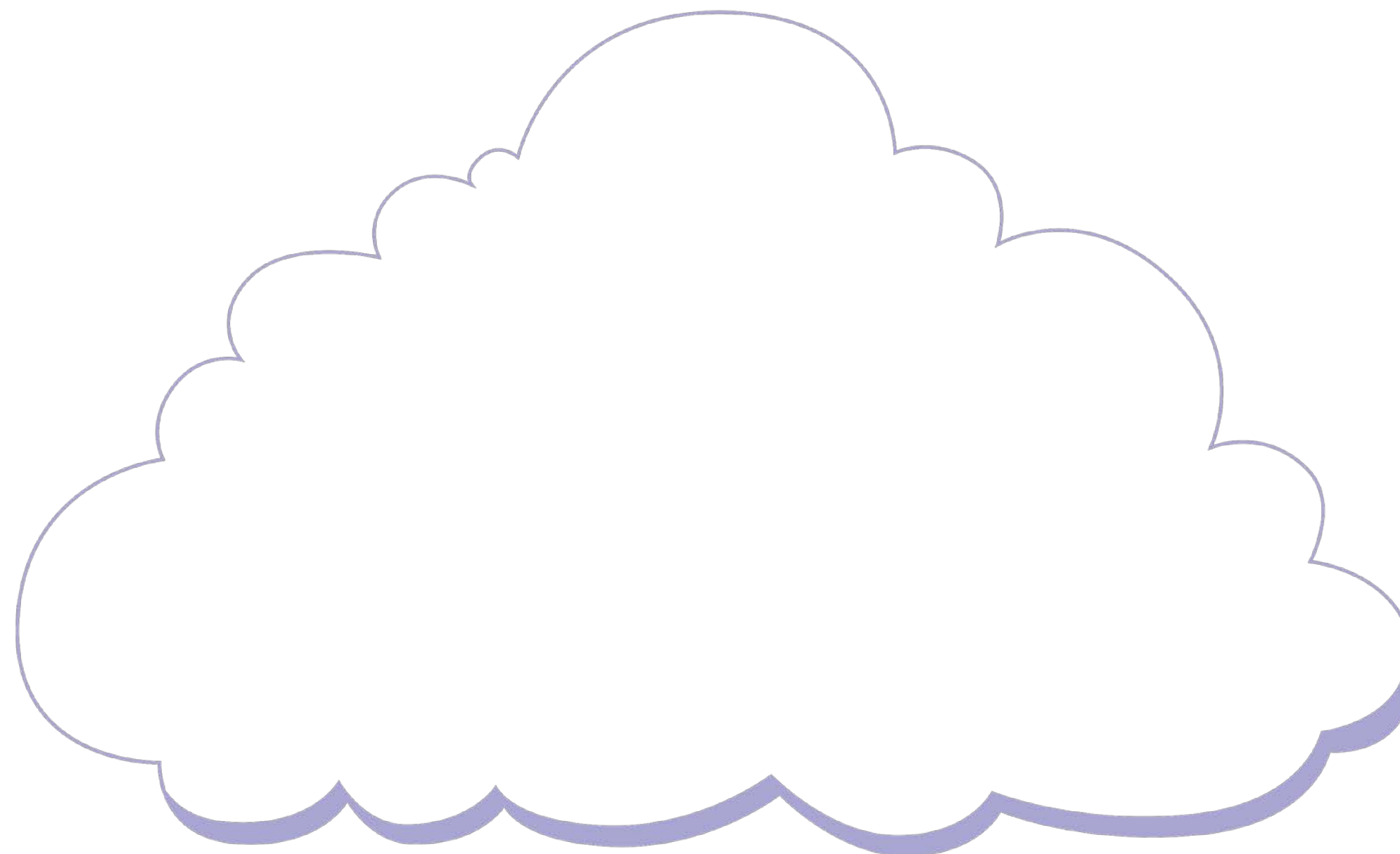
O problema de regressão visa agrupar observações em *clusters* semelhantes, enquanto o problema de classificação visa identificar padrões ou instâncias incomuns nos dados.

O problema de regressão visa prever classes ou categorias, enquanto o problema de classificação visa prever valores contínuos.

O problema de regressão visa identificar padrões ou instâncias incomuns nos dados, enquanto o problema de classificação visa agrupar observações em *clusters* semelhantes.



# Nuvem de palavras



© Getty Images

O que nós  
**aprendemos  
hoje?**



© Getty Images

O que nós  
**aprendemos  
hoje?**

## Então ficamos assim...

- 1** A classificação em Aprendizagem de Máquina ajuda a categorizar dados, como identificar doenças em pacientes com base em sintomas e exames, facilitando decisões médicas mais rápidas e precisas;
- 2** Modelos de Aprendizagem de Máquina são usados por empresas financeiras para detectar fraudes, analisando transações e comportamentos para identificar atividades suspeitas e proteger clientes;
- 3** Empresas utilizam Aprendizagem de Máquina para dividir clientes em grupos por características como compras e preferências, melhorando o direcionamento de marketing e o desenvolvimento de produtos.



# Saiba mais

Você sabe o que é Aprendizagem de Máquina ou Aprendizado de Máquina?

Nesse artigo você vai entender como esses sistemas são treinados usando dados e algoritmos que lhes permitem melhorar seu desempenho para realizar seu objetivo.

Em outras palavras, a Aprendizagem de Máquina permite que computadores aprendam com os dados, sem que seja preciso explicitamente programar essa tarefa.

COSTA, M.; PYRES, L. Direto ao ponto: o que é Machine Learning com exemplos reais. *Alura*, 19 jan. 2024.

Disponível em:

<https://www.alura.com.br/artigos/machine-learning>.

Acesso em: 10 maio 2024.



# Referências da aula

CAMARGO, R; RIBAS, T. *Gestão ágil de projetos*: as melhores soluções para suas necessidades. São Paulo: Saraiva, 2019.

CHIAVENATO, I. *Administração*: teoria, processo e prática. São Paulo: Manole, 2014.

GUESSE, L. Tipos de problemas de Machine Learning. *Medium*, 10 maio 2023. Disponível em: <https://medium.com/@lucasguesse/tipos-de-problemas-de-machine-learning-1517616c66fd>. Acesso em: 10 maio 2024.

Identidade visual: imagens © Getty Images

**Educação  
Profissional  
Paulista**

Técnico em  
**Ciência de  
Dados**