



O que é Machine Learning ou Aprendizado de Máquina?

É uma subárea da inteligência artificial que visa simular a capacidade humana de aprender.

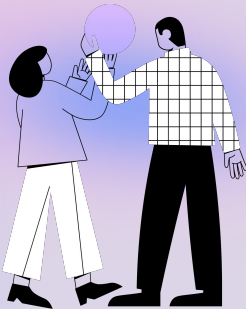
Atualmente, podemos criar modelos que aprendem a fazer **reconhecimento facial**, **predizem se um determinado tumor é maligno ou benigno**, **classificam imagens de animais**, **definem qual seria o valor ideal da venda de um imóvel**, entre diversas outras aplicações.

O processo de aprendizado desses modelos acontece a partir da **associação de diversos dados sobre o problema apresentado**.

Dessa forma, **os modelos aprendem padrões** sobre os dados (*observações*) para conseguir chegar na **resposta esperada (target)**.

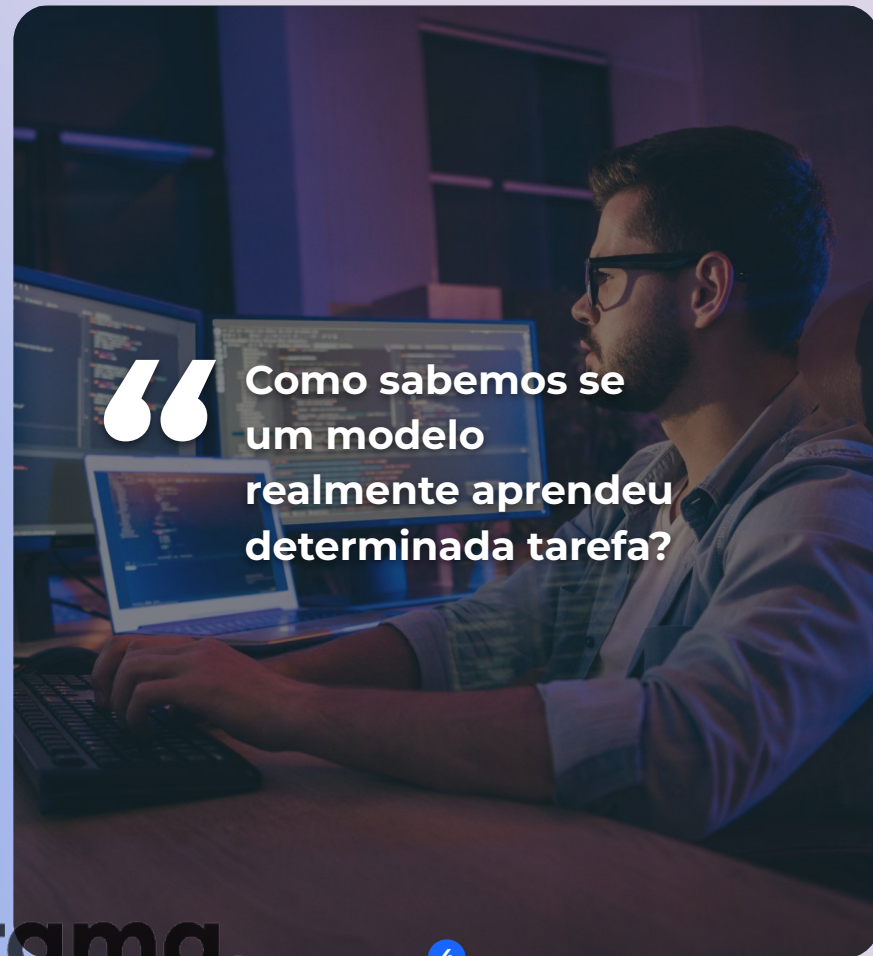
Veja o exemplo a seguir:

Observações	Target
Imagens de animais	Qual animal foi identificado?
Dados sobre o estilo de vida, condição de saúde e alimentação de uma pessoa	Probabilidade da pessoa desenvolver uma doença cardíaca

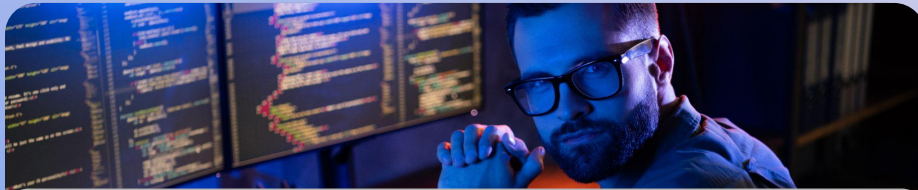


Os problemas de *Machine Learning*:

- **Problemas de classificação:** os modelos devem fazer a previsão de uma classe/categoria para a observação dada. Por exemplo: dada uma imagem de um cachorro, o modelo prediz a raça do animal.
- **Problemas de regressão:** os modelos devem fazer a previsão de um valor numérico (dado contínuo) para a observação colocada. Por exemplo: estipulada uma observação com as características de um imóvel, sua localização e tempo de construção, o modelo prediz qual seria o valor ideal de venda.

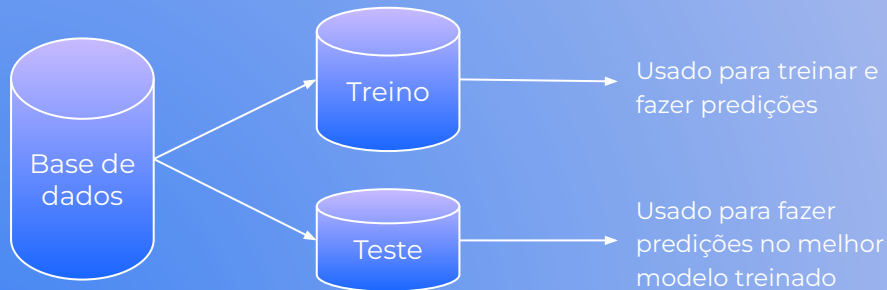


“ Como sabemos se um modelo realmente aprendeu determinada tarefa? ”



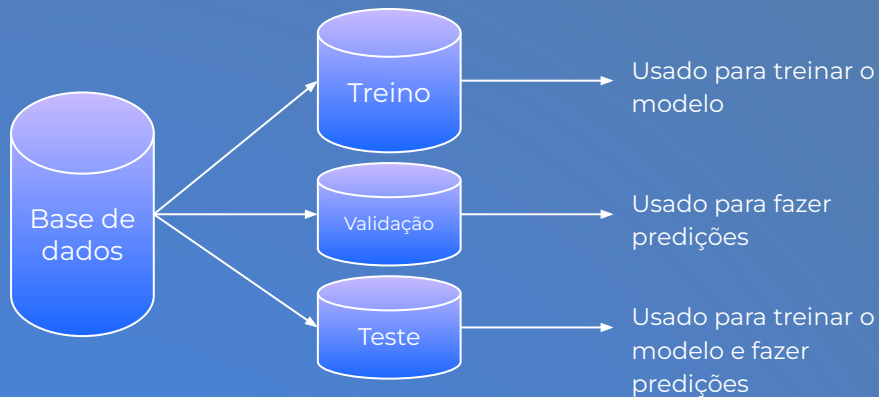
Bom, para sabermos o desempenho de um modelo é preciso avaliá-lo. Para isso, é importante dividirmos a base de dados em **treino e teste** ou em **treino, teste e validação**.

A maior parte dos dados ficam para o treinamento, porque **quanto mais exemplos diferentes o modelo conhecer, mais ele aprende**. Os demais dados ficam para teste ou metade para teste e a outra metade para validação. Como mostra a imagem a seguir:



Os **dados de treino** são usados para **treinar os modelos e fazer previsões**. Com estas previsões feitas é possível ajustar os parâmetros dos modelos e testá-los usando outros algoritmos até que se encontre o melhor modelo possível.

Os **dados de teste** são usados para **realizar previsões** no melhor modelo encontrado para identificarmos a performance do modelo quando ele é exposto a dados que ele não conheceu no processo de treinamento.



Os **dados de treino** são usados para **treinar os modelos**, enquanto que os **dados de validação** são usados para **fazer as previsões**.

Estas **previsões** são usadas para **ajustar os modelos**. Já os **dados de teste** são usados para **realizar previsões** no melhor modelo encontrado.

A **avaliação da performance do modelo** é muito importante. Com ela conseguimos saber se o modelo aprendeu determinada tarefa ou não.

A avaliação pode ser feita de duas maneiras:

1. **Considerando o quanto o modelo está errando;**
2. **Evidenciando o quanto o modelo está acertando.**

Para os **problemas de regressão**, a diferença entre os valores retornados pelo modelo e o valor esperado (o target da base de dados) são calculados.

Para os **problemas de classificação**, normalmente, calcula-se quantas observações foram definidas em classes corretas ou em classes erradas.