

# Aprendizado de máquina

**Machine Learning** é a área da ciência que busca fazer com que computadores ou máquinas sejam capazes de aprender tarefas, sem terem sido explicitamente programados. Aprendizado de máquina ou Machine Learning, da abreviação **ML**, é considerada um tópico de **inteligência artificial** dentre os vários tópicos que compõe o assunto. Dentro do escopo de ML, podemos citar os principais tipos de treinamento: **Algoritmos de aprendizado supervisionado**, algoritmos de aprendizado não supervisionado, algoritmos de aprendizado por Reforço e sistemas de recomendação. O repositório trata especialmente do aprendizado supervisionado.

## 1 Aprendizado Supervisionado

Os algoritmos de aprendizado supervisionado tratam-se de mapeamentos do tipo  $X \rightarrow y$ , onde o símbolo  $X$  é chamado de entrada, características de entrada ou variável de entrada e  $y$  é a saída, rótulo de  $X$  ou variável alvo. A interpretação que  $x$  representa a variável independente e  $y$  depende de  $x$  também é válida.

Alimenta-se o algoritmo com diversas amostras, ou  $(X : \textit{característica}; y : \textit{rótulo})$ . Note que sabemos as saídas corretas para cada entrada, dado um conjunto de amostras ou dados. Em outras palavras, alimenta-se o algoritmo para fazê-lo aprender. Entende-se aprender como a capacidade que o determinado algoritmo possui de encontrar padrões entre  $X$  e  $y$ , ou seja, encontrar o "mapa" que relaciona  $X$  e  $y$ . Após o treinamento através de algoritmos, usando o conjunto de dados que possui as "perguntas" e as "respostas certas", chega-se em um modelo capaz de estimar a resposta para uma "pergunta" ou entrada  $X$  que não faz parte do conjunto de dados de treinamento. Em outras palavras, o modelo pode estimar um valor  $\hat{y}$ , ou dizemos que o modelo atribui um rótulo apropriado para a característica de entrada em questão.

Dentro do tópico aprendizado supervisionado, ainda temos dois subtópicos: **regressão e classificação**. A distinção entre esses dois tipos de algoritmos, está nas possíveis respostas que o modelo aprende a estimar.

Para regressão, a saída pode assumir infinitos valores. Um exemplo de aplicação, consiste no treinamento para dados de tensão e corrente. Seja uma corrente medida, o modelo deve estimar uma tensão para uma entrada que não está no conjunto de dados de treinamento.

Para classificação, a saída pode assumir um conjunto pequeno e finito de valores. Um exemplo de aplicação, consiste no treinamento para dados do tipo tamanho de um tumor e à associação com maligno ou benigno. O modelo deve estimar se é maligno ou benigno para uma entrada que não está no conjunto de dados de treinamento.