



Como funciona o aprendizado supervisionado em classificação

☰ Ciclo	Ciclo 02: Aprendizado supervisionado - Classificação
# Aula	10
🕒 Created	@January 5, 2023 4:41 PM
☑ Done	☑
☑ Ready	☑

Objetivo da Aula:

- ☐ Introdução
- ☐ Aprendizado supervisionado
- ☐ Resumo
- ☐ Próxima aula

Conteúdo:

▼ 1. Introdução

Sistemas de Machine Learning pode ser classificados de acordo com o tipo de aprendizagem.

Os 4 principais tipos de aprendizagem são:

1. Aprendizado Supervisionado
2. Aprendizado Não-Supervisionado

3. Aprendizado Semi-Supervisionado

4. Aprendizado por Reforço

▼ 2. Aprendizado supervisionado

No aprendizado supervisionado, o conjunto de dados para treinamento que você usa para treinar o algoritmo possui características que descrevem o fenômeno observado e um nome ou rótulo que identifica aquela observado e o classifica em classes.

Imagine o algoritmo como uma criança e nós os seus supervisores, como os pais, professores ou educadores. Nós queremos que a criança (computador) aprenda a classificar um observado animal como um cachorro.

Para isso, nós mostramos várias imagens diferentes, algumas são imagens de cachorros, enquanto outras podem ser imagens de qualquer outro animal.

Quando nós mostramos a imagem do cachorro, nós dizemos para a criança: “Esse é um cachorro”. Para as outras imagens podemos dizer que não são um cachorro, sem necessariamente dizer o nome do outro animal em questão.

Após mostrar algumas imagens, nós escolhemos a imagem de um cachorro e perguntamos: “Que animal é esse?”. Se ela responder corretamente, nós a parabenizamos, se ela responder errado, nós a corrigimos: “Não, esse é um cachorro”.

Esse processo de aprendizagem é chamado de aprendizagem supervisionada.

A aprendizagem supervisionada pode resolver 2 problemas clássicos do dia a dia: Classificação e Regressão.

▼ 2.1 Tarefa de Classificação

Quando o conjunto de treinamento possui o rótulo do fenômeno observado como uma variável discreta, dizemos que é um problema de classificação, ou seja, o algoritmo precisa aprender a classificar os exemplos do conjunto, através da interpretação de suas características, em dos rótulos ou classes.

Variáveis discretas são usadas para medir características que podem assumir apenas um número finito contável de valores e, assim, somente fazem sentido valores inteiros. Por exemplo, número de filhos, número de bactérias por litro de leite, número de cigarros fumados por dia e etc.

▼ 2.1.1 Exemplos

1. Classificação de um cliente em bom ou mal pagador.
2. Classificação de uma imagem em gato ou cachorro.
3. Classificação de uma imagem médica em doença ou não-doença.
4. Classificação de uma pessoa para receber ou não um cartão de crédito.
5. Classificação de uma notícia em alguma categoria.
6. Classificação de filme entre 1, 2, 3, 4 ou 5 estrelas.

▼ 2.2 Tarefa de Regressão

Quando o conjunto de treinamento possui o rótulo do fenômeno observado como uma variável real ou contínua, como o salário e o peso, por exemplo.

Nesse tipo de tarefa, o algoritmo tenta ajustar um novo ponto, ao conjunto de pontos do conjunto de dados, de modo que apresente a menor distância possível dos pontos do conjunto de dados. Por exemplo, peso, salário, número de vendas, preço do produto, propensão de compra, número de dias até a primeira venda acontecer, etc.

▼ 2.2.1 Exemplos

1. Previsão de vendas para os próximos 6 meses.
2. Até quanto eu posso aumentar o preço do produto, sem diminuir as vendas.
3. Qual o número médio de dias esperado até a venda do produto acontecer?
4. Qual a probabilidade de um cliente que já possui um produto, comprar o segundo, dentro de 30 dias?
5. Qual a probabilidade dessa imagem ser de um cachorro?

▼ 2.3 Algoritmos de ML com aprendizado supervisionado

▼ 3.2.1 Classificação

1. K-Nearest Neighbors
2. Naive Bayes
3. Light Gradient Boost Machine (LGBM)
4. Categorical Boost (CatBoost)
5. Logistic Regression
6. Support Vector Machine (SVM)
7. Decision Tree
8. Random Forests
9. Neural Networks

▼ 3.2.2 Regressão

1. Linear Regression
2. Polinomial Regression

▼ 3. Resumo

1. Existem 4 tipos de aprendizado em Machine Learning: Supervisionado, Não-supervisionado, Semi-Supervisionado e por Reforço.
2. No aprendizado supervisionado, o conjunto de dados com as características que descrevem o fenômeno observado são rotulados.
3. Algoritmos de aprendizado supervisionado resolvem os problemas de Classificação e Regressão.

▼ 4. Próxima aula

K-Nearest Neighbors - Teoria