



O funcionamento da classificação de dados

Como funciona a classificação de dados



A classificação de dados deve ser utilizada sempre que estiver disponível uma base de dados para ser a fonte de conhecimento de criação de um modelo, que poderá ser utilizado posteriormente para a solução de novas situações.

O que é preciso para classificar?

- Base de dados com ocorrências e suas respectivas classes;
- Algoritmo que irá aprender com os dados e criar o modelo;
- O processo de classificação acontece em duas etapas: treinamento e aplicação.

Amostra de base de dados com diagnósticos conhecidos

Tabela 1 - Cadastro de pacientes e seus respectivos diagnósticos

Idade	Sexo	Peso	Altura	Estado Civil	Profissão	Jornada	Intervalo	Horas	Dieta	Atividade física	Diagnóstico
12	F	40	1,58	Solteiro	Estudante	0	Sim	2	?	Sim	Cifose
15	F	50	1,57	Solteiro	Estudante	0	Sim	2	Não	Não	Escoliose
32	F	71	1,65	Casado	Auxiliar de limpeza	8	Sim	1	?	?	Escoliose
14	M	57	1,72	Solteiro	Estudante	0	Sim	2	?	?	Escoliose
31	M	61	1,71	Casado	Técnico eletrônico	8	Sim	1	Não	Sim	Cervicobraquialgia
36	M	90	1,83	Solteiro	Chaveiro	11	Não	0	Não	Sim	Fratura
20	M	78	1,69	Solteiro	Estudante	0	Sim	2	?	Sim	Lesão
50	M	88	1,70	Casado	Motorista	12	Sim	2	Não	Sim	Lesão
43	F	58	1,70	Divorciado	Representante comercial	44	Sim	2	?	?	Lesão
31	M	89	1,84	Casado	Representante comercial	8	Sim	1	Sim	Sim	Lesão
22	M	85	1,75	Solteiro	Estudante	0	Sim	2	Não	Sim	Lesão
45	M	78	1,76	Casado	Técnico em manutenção	?	Sim	2	Não	Não	Hérnia de disco
24	M	79	1,70	?	Atleta	?	Sim	2	?	Sim	Osteíte púbica
76	F	63	1,58	Viúvo	Do lar	44	Não	0	Não	Não	Tendinose
70	M	88	1,74	Casado	Professor	4	Sim	2	Não	Não	Ruptura de tendão do quadríceps
11	M	60	1,59	Solteiro	Estudante	0	Sim	2	Não	Sim	Fratura

Fonte: Dados da pesquisa.

Algoritmos para o treinamento do modelo

- Árvores de decisão;
- Redes neurais;
- Métodos estatísticos;
- Formulações híbridas.

O treinamento

Durante o treinamento o algoritmo utiliza o conjunto de dados oferecido para identificar quais características apresentam os indivíduos com cada um dos diagnósticos apresentados.

- Cada um dos conjuntos de características e diagnóstico apresentado é chamado de padrão de entrada.
- As características do padrão de entrada são chamadas de atributos;
- O diagnóstico é chamado de rótulo.

O que o algoritmo aprenderá?

O algoritmo aprenderá como classificar ou diagnosticar:

- Cifose;
- Escoliose;
- Cervicobraquialgia;
- Fratura;
- Lesão;
- Hérnia de disco;
- Osteíte púbica;
- Tendinose;
- Ruptura de tendão do quadríceps.

O funcionamento da classificação de dados

O que isso permitirá?

Ao chegar um novo indivíduo suas características podem ser informadas ao modelo criado, mesmo sem ter tido contato com um paciente, o modelo será capaz de determinar qual dos diagnósticos que ele conhece é adequado para o indivíduo.

Avaliação do modelo

Antes de ser colocado na prática é importante que os modelos, além de serem treinados, sejam validados. O ideal é que os resultados que ele produz sejam avaliados com dados diferentes do utilizado para treinamento ou com a presença de um profissional especializado no assunto.



Obrigada!

Ana Laurentino