

05

MultiLabelBinarizer

Ao analisar os produtos comprados por clientes em uma loja, temos para cada cliente e data, uma lista do Python com os produtos que foram comprados.

Para passar a informação de uma forma que o modelo possa interpretar, podemos criar uma coluna para cada produto, e utilizar um valor binário(0,1) para identificar se aquele produto foi comprado em determinada linha:

DATA	CLIENTE	PRODUTOS
24/07/2019	ANA	[MACARRÃO, TOMATE]
24/07/2019	ANA	[TRIGO, CARNE]
22/07/2019	PEDRO	[CARNE, TOMATE]
23/07/2019	ANA	[OVOS, CEBOLA, ALHO]
25/07/2019	BRUNO	[CEBOLA, ALHO]

Considerando a tabela acima, após o tratamento com MultiLabelBinarizer, quantos novas colunas teremos?

Seleciona uma alternativa

A Sete colunas.

B Nove colunas.

C Cinco colunas.

D Onze colunas.