

Fundamentos de OneHotEncoding

Atributo	A	B	C
A	1	0	0
B	0	1	0
A	1	0	0
C	0	0	1

OneHotEncoding

- OneHotEncoding produz um único atributo de saída com uma matriz densa, a partir de n atributos numéricos
- Espera atributos numéricos: Podemos usar StringIndexer para transformar

Geography	Gender
France	Female
Spain	Female
France	Female
France	Female
Spain	Female
Spain	Male
France	Male
Germany	Female
France	Male
France	Male
France	Male
Spain	Male
France	Female
France	Female
Spain	Female
Germany	Male
Germany	Male
Spain	Female
Spain	Male
France	Female

indexer_c1	indexer_c2
0.0	1.0
2.0	1.0
0.0	1.0
0.0	1.0
2.0	1.0
2.0	0.0
0.0	0.0
1.0	1.0
0.0	0.0
0.0	0.0
0.0	0.0
2.0	0.0
0.0	1.0
0.0	1.0
2.0	1.0
1.0	0.0
1.0	0.0
2.0	1.0
2.0	0.0
0.0	1.0

onehot_c1	oneho_c2
(2,[0],[1.0])	(1,[],[])
(2,[],[])	(1,[],[])
(2,[0],[1.0])	(1,[],[])
(2,[0],[1.0])	(1,[],[])
(2,[],[])	(1,[],[])
(2,[],[])	(1,[0],[1.0])
(2,[0],[1.0])	(1,[0],[1.0])
(2,[1],[1.0])	(1,[],[])
(2,[0],[1.0])	(1,[0],[1.0])
(2,[0],[1.0])	(1,[0],[1.0])
(2,[0],[1.0])	(1,[0],[1.0])
(2,[],[])	(1,[0],[1.0])
(2,[0],[1.0])	(1,[],[])
(2,[0],[1.0])	(1,[],[])
(2,[],[])	(1,[],[])
(2,[1],[1.0])	(1,[0],[1.0])
(2,[1],[1.0])	(1,[0],[1.0])
(2,[],[])	(1,[],[])
(2,[],[])	(1,[0],[1.0])
(2,[0],[1.0])	(1,[],[])