



UNIVERSIDADE DO MINHO
DEPARTAMENTO DE INFORMÁTICA
DESCOBERTA DE CONHECIMENTO

Ficha 7

5 de Abril de 2019

Francisco Oliveira (A78416)



Conteúdo

1 Resolução de exercícios

2

1 Resolução de exercícios

A primeira fase na construção de um modelo para obter uma árvore de decisão é o tratamento dos dados. Para isso começamos a selecionar os dados necessários para obter resultados mais eficientes. Daí aplicou-se um filtro no dataset *Titanic training*, com o operador *Select Attributes*, para se utilizar apenas os valores não vazios. Já em relação ao *Titanic Test*, apenas selecionamos os atributos relevantes, como se mostra na imagem abaixo, descartando a classe *Survived*, pois é esta que queremos descobrir na classificação.

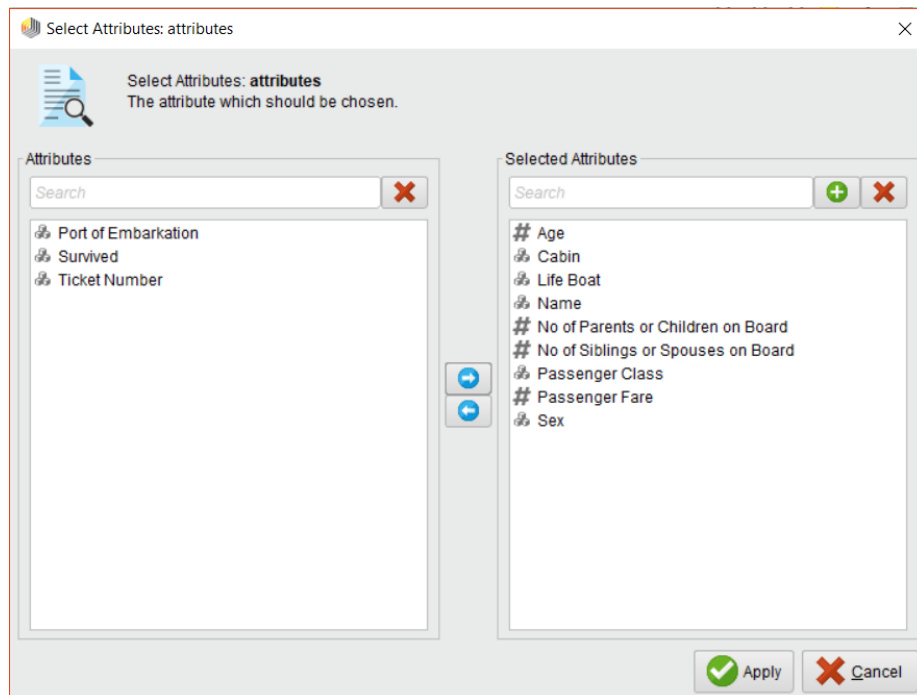


Figura 1: Seleção de atributos

De seguida, aplicamos o operador *Set Role*, para indicar que o atributo *Survived* é a classe no dataset *Titanic Train* e indicar que o atributo Nome é o identificador no dataset *Titanic Test*. Por fim, implementaram-se os operadores *Decision Tree* e *Apply Model*, para obter os resultados finais.

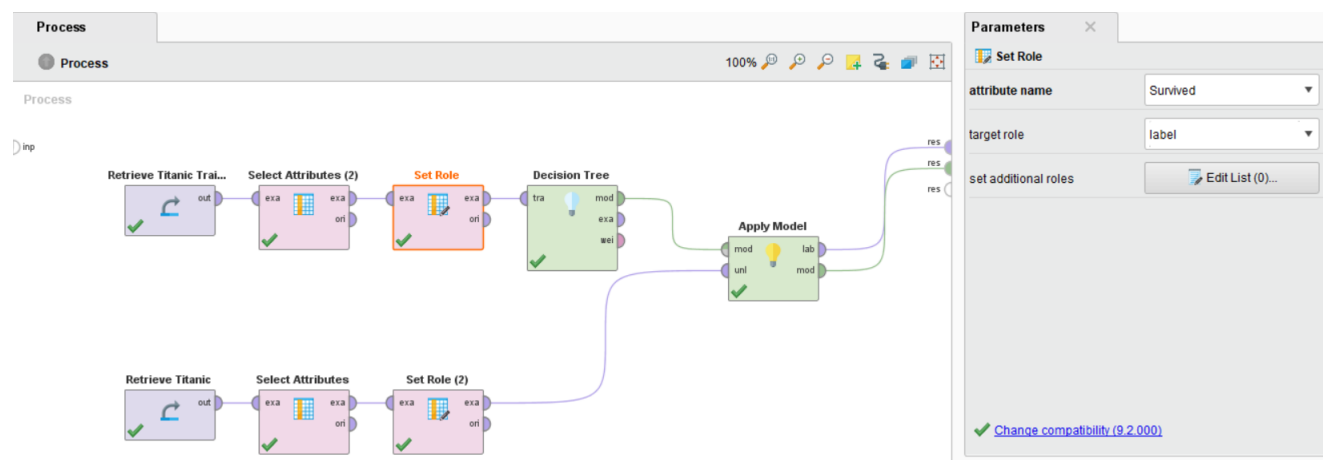


Figura 2: Modelo de árvore de decisão

Gerou-se uma árvore com profundidade máxima de 7. A conclusão de maior relevância que se pode extrair da árvore de decisão é que o sexo da pessoa teve uma grande influência na sobrevivência dos passageiros do Titanic, sendo que pessoas do sexo feminino tiveram uma maior probabilidade de ter sobrevivido. A idade também teve um papel importante quanto à sobrevivência, pelo que as crianças tiveram prioridade em ser salvas.

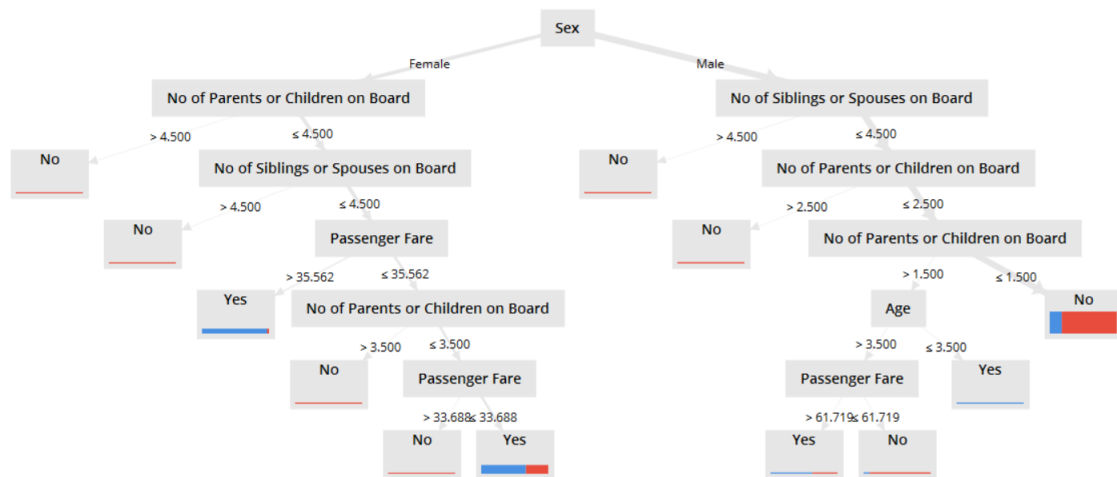


Figura 3: Árvore de decisão

Eis a árvore de decisão traduzida numa tabela:

Row No.	Name	prediction(S...	confidence(...	confidence(...	Passenger ...	Sex	Age	No of Sibling...	No of Parent...	Passenger F...	Cabin	Life Boat
1	Allen, Miss. E...	Yes	0.971	0.029	First	Female	29	0	0	211.338	B5	2
2	Allison, Mast...	Yes	1	0	First	Male	0.917	1	2	151.550	C22 C26	11
3	Allison, Miss. ...	Yes	0.971	0.029	First	Female	2	1	2	151.550	C22 C26	?
4	Allison, Mr. H...	Yes	0.625	0.375	First	Male	30	1	2	151.550	C22 C26	?
5	Allison, Mrs. ...	Yes	0.971	0.029	First	Female	25	1	2	151.550	C22 C26	?
6	Anderson, Mr...	No	0.181	0.819	First	Male	48	0	0	26.550	E12	3
7	Andrews, Mis...	Yes	0.971	0.029	First	Female	63	1	0	77.958	D7	10
8	Andrews, Mr. ...	No	0.181	0.819	First	Male	39	0	0	0	A36	?
9	Appleton, Mrs...	Yes	0.971	0.029	First	Female	53	2	0	51.479	C101	D
10	Artagaveytia, ...	No	0.181	0.819	First	Male	71	0	0	49.504	?	?
11	Astor, Col. Jo...	No	0.181	0.819	First	Male	47	1	0	227.525	C62 C64	?
12	Astor, Mrs. Jo...	Yes	0.971	0.029	First	Female	18	1	0	227.525	C62 C64	4
13	Aubart, Mme. ...	Yes	0.971	0.029	First	Female	24	0	0	69.300	B35	9
14	Barber, Miss. ...	Yes	0.971	0.029	First	Female	26	0	0	78.850	?	6
15	Barkworth, Mr...	No	0.181	0.819	First	Male	80	0	0	30	A23	B
16	Baumann, Mr...	No	0.181	0.819	First	Male	?	0	0	25.925	?	?
17	Baxter, Mr. Qu...	No	0.181	0.819	First	Male	24	0	1	247.521	B58 B60	?
18	Baxter, Mrs. J...	Yes	0.971	0.029	First	Female	50	0	1	247.521	B58 B60	6

Figura 4: Tabela de resultados