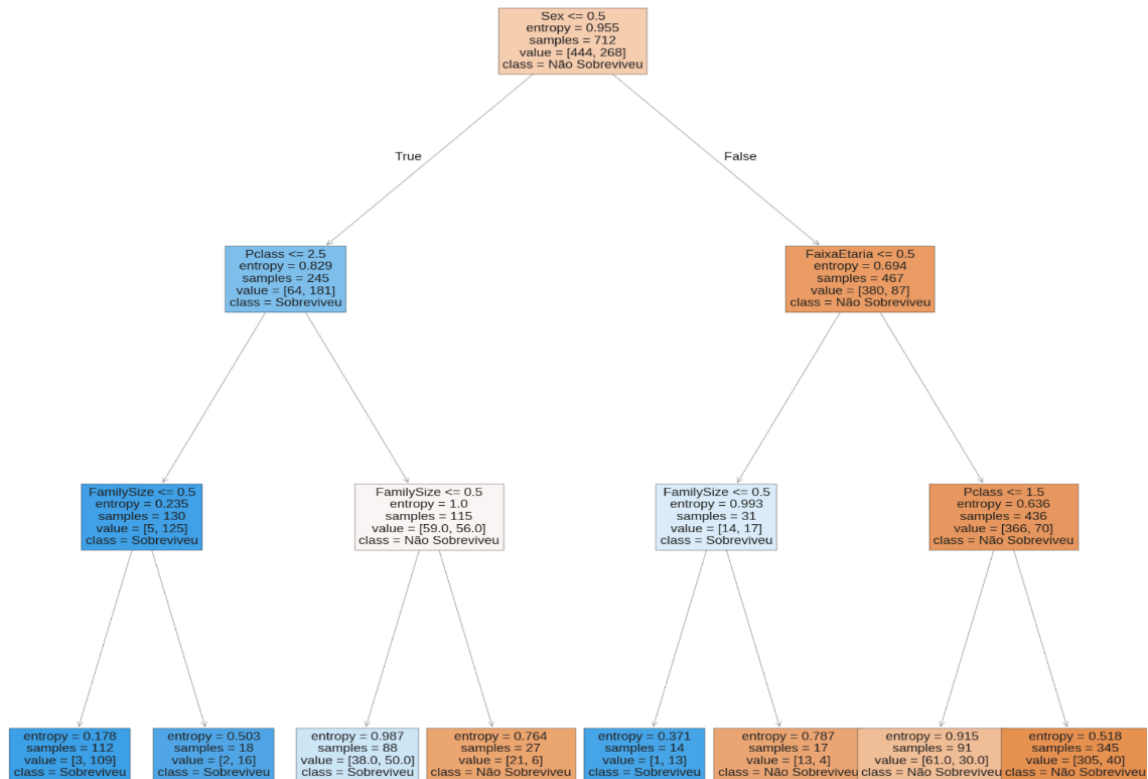


// 801779
// Giuseppe Cordeiro
// lista 2 IA

- 1) A partir do resultado gerado pela árvore de decisões, podemos notar que a maior parte dos sobreviventes do titanic eram mulheres que estavam localizadas na primeira e segunda classe do Titanic.



- 2)
- 1- O algoritmo ID3 utiliza o ganho de informação simples para selecionar atributos, enquanto o C4.5 aprimora essa abordagem ao empregar a razão de ganho, o que ajuda a evitar o viés em relação a atributos com muitos valores. Além disso, o ID3 trabalha apenas com atributos discretos e não lida com valores ausentes. Já o C4.5 é mais versátil, permitindo o uso de atributos contínuos, tratando dados faltantes por meio de probabilidades e realizando a poda da árvore para reduzir o overfitting
 - 2- O C4.5 organiza os valores dos atributos numéricos em ordem e testa diversos pontos de corte. Para cada ponto, ele avalia o ganho (ou *gain ratio*) obtido ao dividir os dados em dois grupos: um com valores abaixo do limiar e outro com valores acima. O ponto que proporciona o maior benefício é escolhido, convertendo o atributo contínuo em uma variável binária de divisão.

- 3) c) Iris_Versicolor, Iris_Setosa, Iris_Versicolor, Iris_Virginica

4) c) I e II, apenas.

5)

	Precisão	Recall	F1 Score	TVP	TFN	TFP	TVN
A	0,588	0,588	0,588	10/17	7/17	7/105	98/105
B	0,652	0,833	0,731	15/18	3/18	8/104	96/104
C	0,769	0,667	0,714	20/30	10/30	6/92	86/92
D	0,893	0,877	0,885	50/57	7/57	6/65	59/65