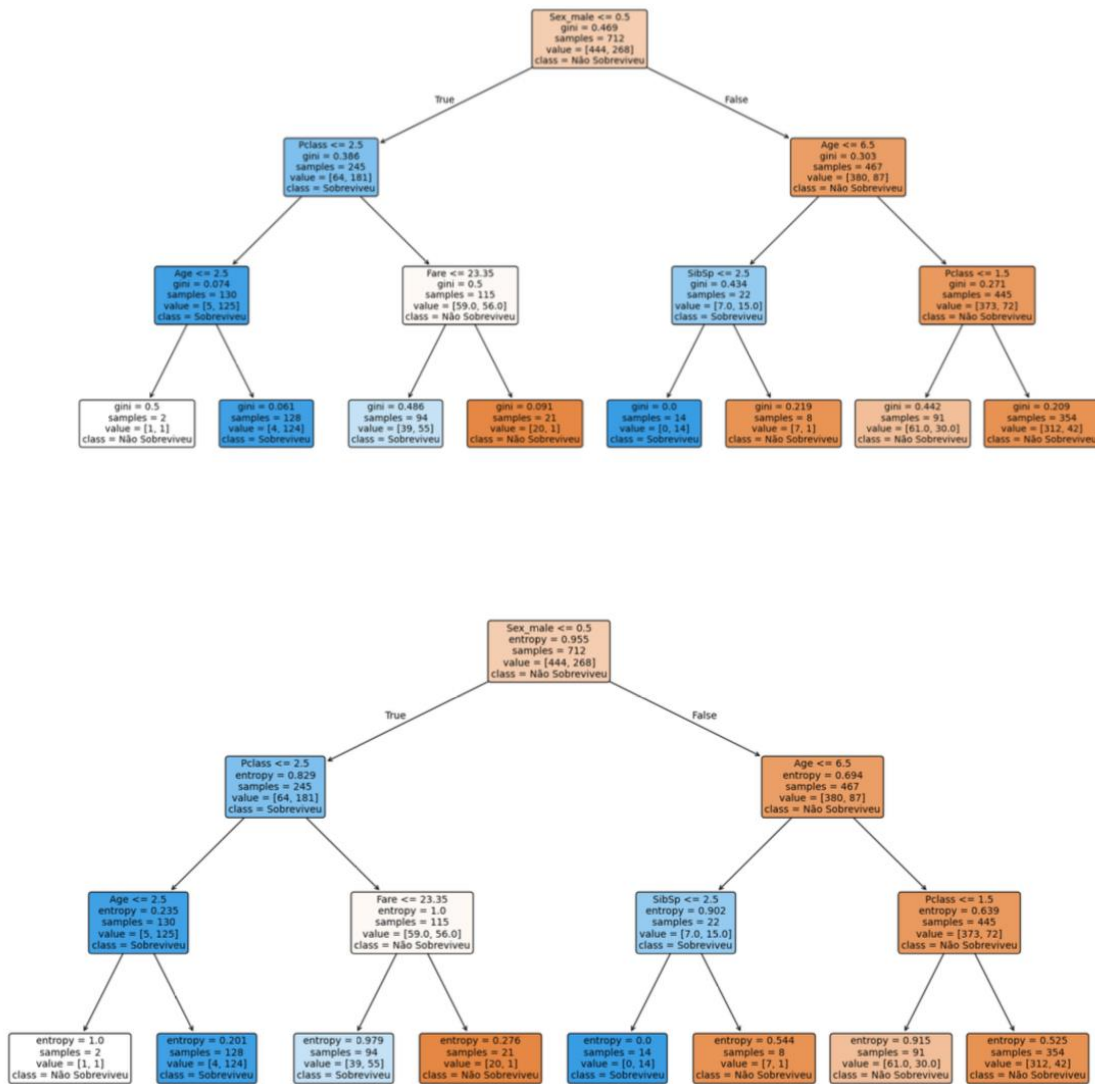


Lista IA

Questão 1: 1. A acurácia obtida foi de 0.7821 com Gini e 0.7933 com Entropia. A estrutura das árvores apresenta diferenças internas, embora visualmente possam parecer similares.



Questão 2: 2. Parâmetros como profundidade máxima, número mínimo de amostras por folha e critério de divisão afetam o desempenho da árvore. Um modelo muito profundo pode sofrer de overfitting, enquanto um muito raso pode não capturar bem os padrões dos dados.

Questão 3: 3. O GridSearchCV explora todas as combinações possíveis de hiperparâmetros, o RandomizedSearchCV faz uma busca aleatória dentro de um espaço definido, e o BayesSearchCV usa otimização bayesiana para escolher as melhores combinações de forma eficiente.

Questão 4: 4. A alternativa correta é A. A) Apenas I e II

Questão 5: 5. A resposta correta é A. A) Apenas I e II

Questão 6: 6. O C4.5 melhora o ID3 ao suportar atributos contínuos, lidar com valores ausentes, usar gain ratio para escolha de atributos, permitir poda da árvore e gerar regras a partir da árvore, tornando-o mais robusto e preciso.

Questão 7: 7. O Ganho de Informação (Information Gain) mede a redução da entropia ao dividir os dados com base em um atributo. Ele favorece atributos com muitos valores distintos, o que pode levar a divisões ruins. A Razão de Ganho (Gain Ratio) corrige esse viés normalizando o Ganho de Informação pelo Ganho Intrínseco, que mede a diversidade dos valores do atributo. Assim, ela evita escolher atributos apenas por terem muitos valores únicos e seleciona divisões mais equilibradas.