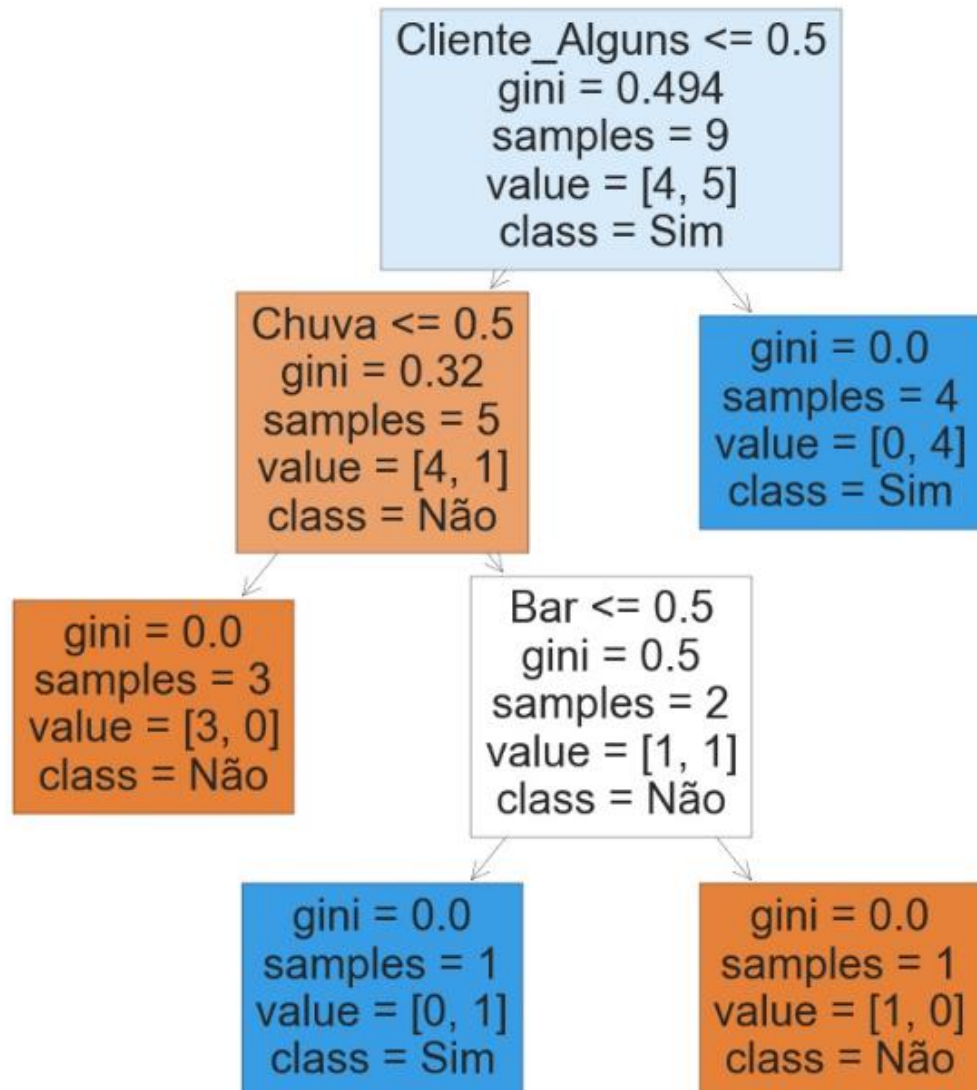


Ciência da Computação – ICEI PUC Minas Praça da Liberdade

Aluno: Leon Júnio Martins Ferreira

INTELIGÊNCIA ARTIFICIAL

1. A árvore plotada:



2. Não houve ganho algum, manteve a mesma accuracy e precisão. Mas a árvore resultante ficou diferente nos casos, visto que a raiz mudou de clientes para tempo.
3. Os hiperparâmetros descobertos:

```
{'criterion': 'gini', 'max_depth': None, 'max_features': None, 'min_samples_split': 2,  
'min_weight_fraction_leaf': 0.2, 'random_state': 0, 'splitter': 'best'}
```

Score: 0.8888888888888888

A classificação após os hiperparâmetros:

```
print(classification_report(y_teste, previsoes))
```

	precision	recall	f1-score	support
Não	0.67	1.00	0.80	2
Sim	0.00	0.00	0.00	1
accuracy			0.67	3
macro avg	0.33	0.50	0.40	3
weighted avg	0.44	0.67	0.53	3

4. Tanto o algoritmo de árvore ID3 quanto o algoritmo C4.5 são algoritmos de aprendizado de máquina supervisionado que são usados para criar modelos de árvore de decisão. Segundo o artigo que foi lido, os dois algoritmos têm uma estrutura bem semelhante de árvore em que cada nó da árvore representa um teste de uma variável e cada aresta da árvore representa o resultado de um teste. Porém apresentam diferenças na manipulação de alguns atributos tal como, o ID3 só trabalha atributos categóricos discretos que não tem relação entre si enquanto o C4.5 trabalha com ambos numéricos e nominais/categóricos. O c4.5 tem diferenças para podar a árvore, ele permite que seja podada a árvore para evitar overfitting, e ele usa uma técnica chamada ganho de informação relativa, que leva em consideração o número de valores possíveis de um atributo para evitar a escolha de atributos com muitos valores, o que pode levar à overfitting. Em resumo, o algoritmo C4.5 é uma extensão do ID3 que possui recursos adicionais para lidar com atributos contínuos, valores faltantes, seleção de atributos e poda da árvore.
5. O algoritmo C4.5 lida com atributos numéricos através de um processo de discretização, que divide o intervalo de valores em faixas discretas com base em sua entropia de classe. Depois de determinar os intervalos discretos, o algoritmo usa-os como se fossem atributos categóricos em uma árvore de decisão. Em seguida, o algoritmo segue seu procedimento normal para construir a árvore de decisão, selecionando o atributo que fornece o maior ganho de informação em cada etapa da construção da árvore.