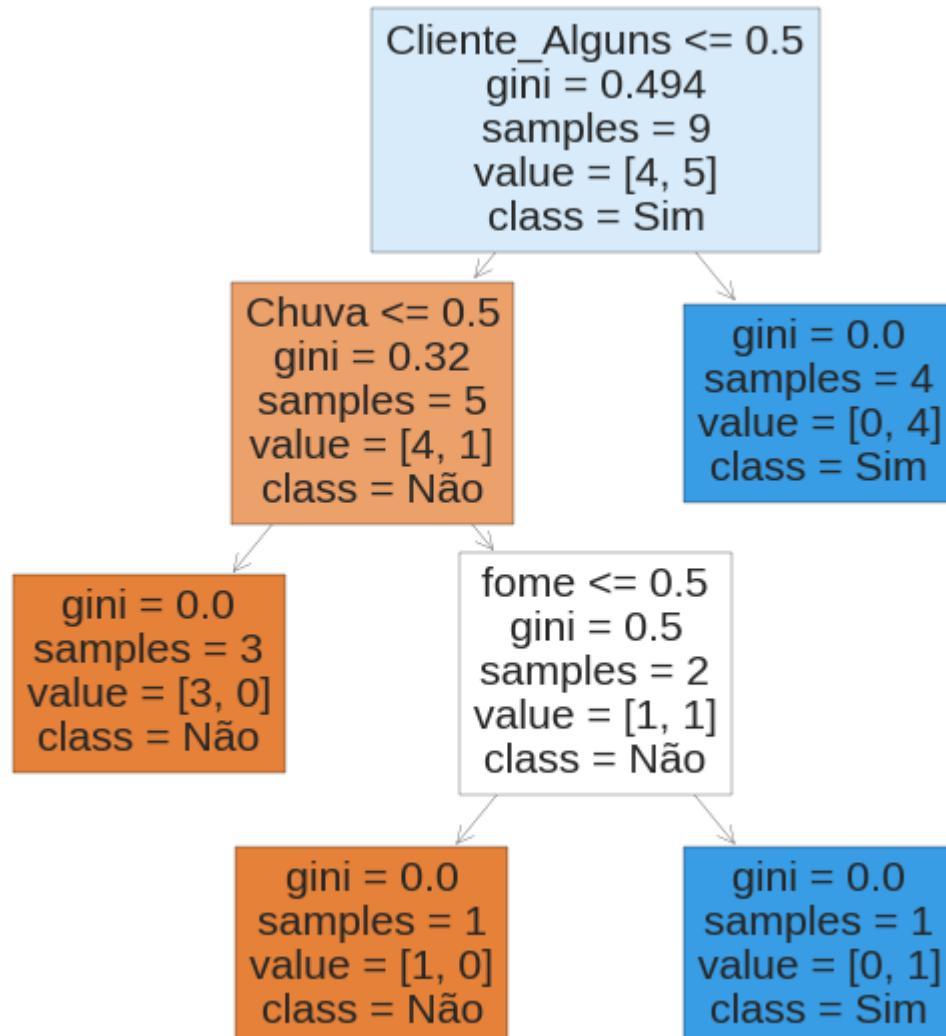


Lista II - Inteligência Artificial

Aluno: Vinicius Gabriel dos Santos Teixeira

Questão 1

1)



2) Não houve ganho, possivelmente pois a base de dados é muito pequena.

3)

| | |
|--------------------------|------|
| criterion | gini |
| max_depth | None |
| max_features | None |
| min_samples_split | 2 |
| min_weight_fraction_leaf | 0.2 |
| random_state | 0 |
| splitter | best |

Questão 2

- 1) - O ID3 não lida com dados vazios já o C4.5 utiliza média ou moda para completá-los.
- O C4.5 avalia a distribuição dos valores de cada atributo favorecendo atributos que não possuam tendências de apenas um valor.
- O C4.5 consegue lidar com atributos com valores contínuos dividindo-os em pequenas faixas, algo que o ID3 não consegue.
- O C4.5 possui poda de nodos já o ID3 não.
Em geral o C4.5 é um algoritmo mais preciso que o ID3 porém demora mais tempo para completar sua execução.
- 2) O algoritmo C4.5 lida com atributos numéricos utilizando a discretização e a regressão. A discretização consiste em dividir o atributo em faixas de valores criando então um atributo categórico com tais faixas, já a regressão usa cada valor real para prever a classe criando uma árvore de regressão que estende a árvore de decisão.