

2.

Registro	Tem casa própria	Estado Civil	Rendimentos	Bom Pagador
1	Sim	Solteiro	Alto	Não
2	Não	Casado	Médio	Não
3	Não	Solteiro	Baixo	Não
4	Sim	Casado	Alto	Não
5	Não	Divorciado	Médio	Sim
6	Não	Casado	Baixo	Não
7	Sim	Divorciado	Alto	Sim
8	Não	Solteiro	Médio	Sim
9	Não	Casado	Baixo	Não
10	Não	Solteiro	Médio	Sim
11	Sim	Divorciado	Médio	Não
12	Não	Divorciado	Alto	?

1ª Passo: Calcular o índice de Gini para cada evidência.

2ª Passo: Construir a árvore de decisão.

3ª Passo: Repetir 1 e 2 até que  $G_{ini} = 0$ .

$$G_{ini}(E) = 1 - \sum_{k=1}^m p_k^2 \quad m = \text{número de classes} \\ p_k = \text{proporção de amostras em } p_k$$

# 1ª iteração:

1ª Passo:

$$G_{ini}(\text{Tem Casa Própria}) = \frac{4}{11} \left( 1 - \left( \frac{1}{4} \right)^2 - \left( \frac{3}{4} \right)^2 \right) + \frac{7}{11} \left( 1 - \left( \frac{3}{7} \right)^2 - \left( \frac{4}{7} \right)^2 \right) = 0,448$$

← tem casa própria → bom pagador      ← não tem casa própria → bom pagador  
 ↳ não bom pagador

$$G_{ini}(\text{Estado Civil: Solteiro}) = \frac{4}{11} \left( 1 - \left( \frac{2}{4} \right)^2 - \left( \frac{2}{4} \right)^2 \right) + \frac{7}{11} \left( 1 - \left( \frac{2}{7} \right)^2 - \left( \frac{5}{7} \right)^2 \right) = 0,442$$

$$G_{ini}(\text{Estado Civil: Casado}) = \frac{4}{11} \left( 1 - \left( \frac{0}{4} \right)^2 - \left( \frac{4}{4} \right)^2 \right) + \frac{7}{11} \left( 1 - \left( \frac{4}{7} \right)^2 - \left( \frac{3}{7} \right)^2 \right) = 0,312$$

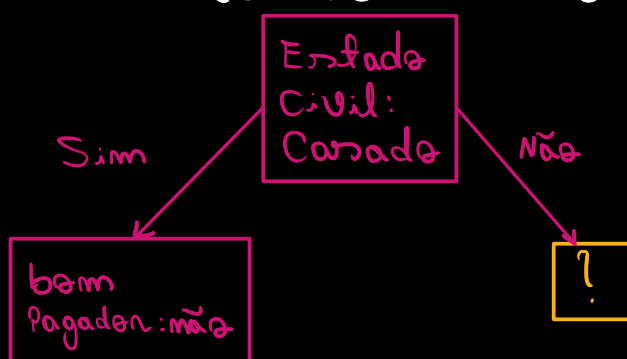
$$G_{ini}(\text{Estado Civil: Divorciado}) = \frac{3}{11} \left( 1 - \left( \frac{2}{3} \right)^2 - \left( \frac{1}{3} \right)^2 \right) + \frac{8}{11} \left( 1 - \left( \frac{2}{8} \right)^2 - \left( \frac{6}{8} \right)^2 \right) = 0,394$$

$$G_{ini}(\text{Rendimentos: Alto}) = \frac{3}{11} \left( 1 - \left( \frac{1}{3} \right)^2 - \left( \frac{2}{3} \right)^2 \right) + \frac{8}{11} \left( 1 - \left( \frac{3}{8} \right)^2 - \left( \frac{5}{8} \right)^2 \right) = 0,462$$

$$G_{ini}(\text{Rendimentos: Médio}) = \frac{5}{11} \left( 1 - \left( \frac{3}{5} \right)^2 - \left( \frac{2}{5} \right)^2 \right) + \frac{6}{11} \left( 1 - \left( \frac{1}{6} \right)^2 - \left( \frac{5}{6} \right)^2 \right) = 0,420$$

$$G_{ini}(\text{Rendimentos: Baixo}) = \frac{3}{11} \left( 1 - \left( \frac{0}{3} \right)^2 - \left( \frac{3}{3} \right)^2 \right) + \frac{8}{11} \left( 1 - \left( \frac{4}{8} \right)^2 - \left( \frac{4}{8} \right)^2 \right) = 0,364$$

2ª Passo: O menor índice Gini é o de Estado Civil: Casado



## # 2ª iteração:

Registro	Tem casa própria	Estado Civil	Rendimentos	Bom Pagador
1	Sim	Solteiro	Alto	Não
2	Não	Casado	Médio	Não
3	Não	Solteiro	Baixo	Não
4	Sim	Casado	Alto	Não
5	Não	Divorciado	Médio	Sim
6	Não	Casado	Baixo	Não
7	Sim	Divorciado	Alto	Sim
8	Não	Solteiro	Médio	Sim
9	Não	Casado	Baixo	Não
10	Não	Solteiro	Médio	Sim
11	Sim	Divorciado	Médio	Não
12	Não	Divorciado	Alto	?

Desconsiderar amostras  
com estado civil: Casado

### 1ª Parte:

$$Gini(\text{Tem casa própria}) = \frac{3}{7} \left( 1 - \left( \frac{1}{3} \right)^2 - \left( \frac{2}{3} \right)^2 \right) + \frac{4}{7} \left( 1 - \left( \frac{3}{4} \right)^2 - \left( \frac{1}{4} \right)^2 \right) = 0,404$$

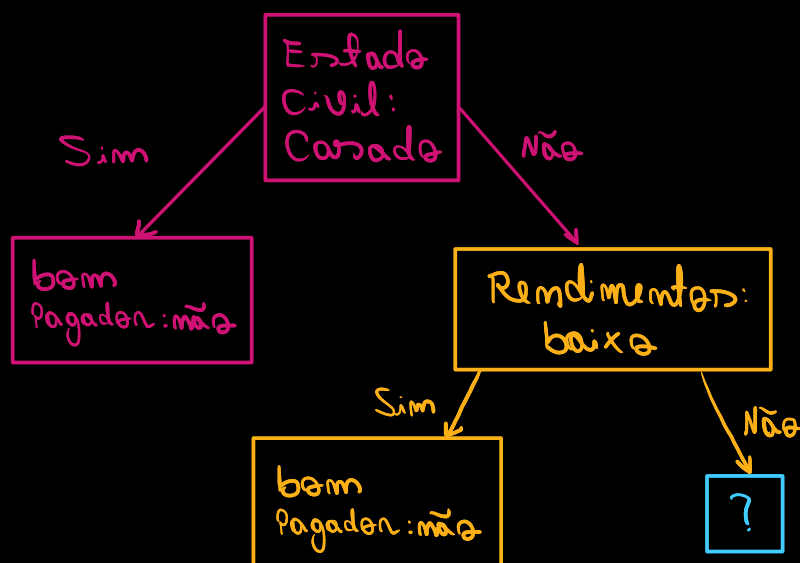
$$Gini(\text{Est. Civil: Solteiro | Divorciado}) = \frac{4}{7} \left( 1 - \left( \frac{2}{4} \right)^2 - \left( \frac{2}{4} \right)^2 \right) + \frac{3}{7} \left( 1 - \left( \frac{2}{3} \right)^2 - \left( \frac{1}{3} \right)^2 \right) = 0,476$$

$$Gini(\text{Rendimentos: Alto}) = \frac{2}{7} \left( 1 - \left( \frac{1}{2} \right)^2 - \left( \frac{1}{2} \right)^2 \right) + \frac{5}{7} \left( 1 - \left( \frac{3}{5} \right)^2 - \left( \frac{2}{5} \right)^2 \right) = 0,486$$

$$Gini(\text{Rendimentos: Médio}) = \frac{4}{7} \left( 1 - \left( \frac{3}{4} \right)^2 - \left( \frac{1}{4} \right)^2 \right) + \frac{3}{7} \left( 1 - \left( \frac{1}{3} \right)^2 - \left( \frac{2}{3} \right)^2 \right) = 0,405$$

$$Gini(\text{Rendimentos: baixo}) = \frac{1}{7} \left( 1 - \left( \frac{0}{1} \right)^2 - \left( \frac{1}{1} \right)^2 \right) + \frac{6}{7} \left( 1 - \left( \frac{4}{6} \right)^2 - \left( \frac{2}{6} \right)^2 \right) = 0,381$$

2ª Parte: O menor índice Gini é o de Rendimentos: baixo



### # 3ª iteração:

Registro	Tem casa própria	Estado Civil	Rendimentos	Bom Pagador
1	Sim	Solteiro	Alto	Não
2	Não	Casado	Médio	Não
3	Não	Solteiro	Baixo	Não
4	Sim	Casado	Alto	Não
5	Não	Divorciado	Médio	Sim
6	Não	Casado	Baixo	Não
7	Sim	Divorciado	Alto	Sim
8	Não	Solteiro	Médio	Sim
9	Não	Casado	Baixo	Não
10	Não	Solteiro	Médio	Sim
11	Sim	Divorciado	Médio	Não
12	Não	Divorciado	Alto	?

Desconsiderar amostras  
com rendimentos: baixa

#### 1ª Parte:

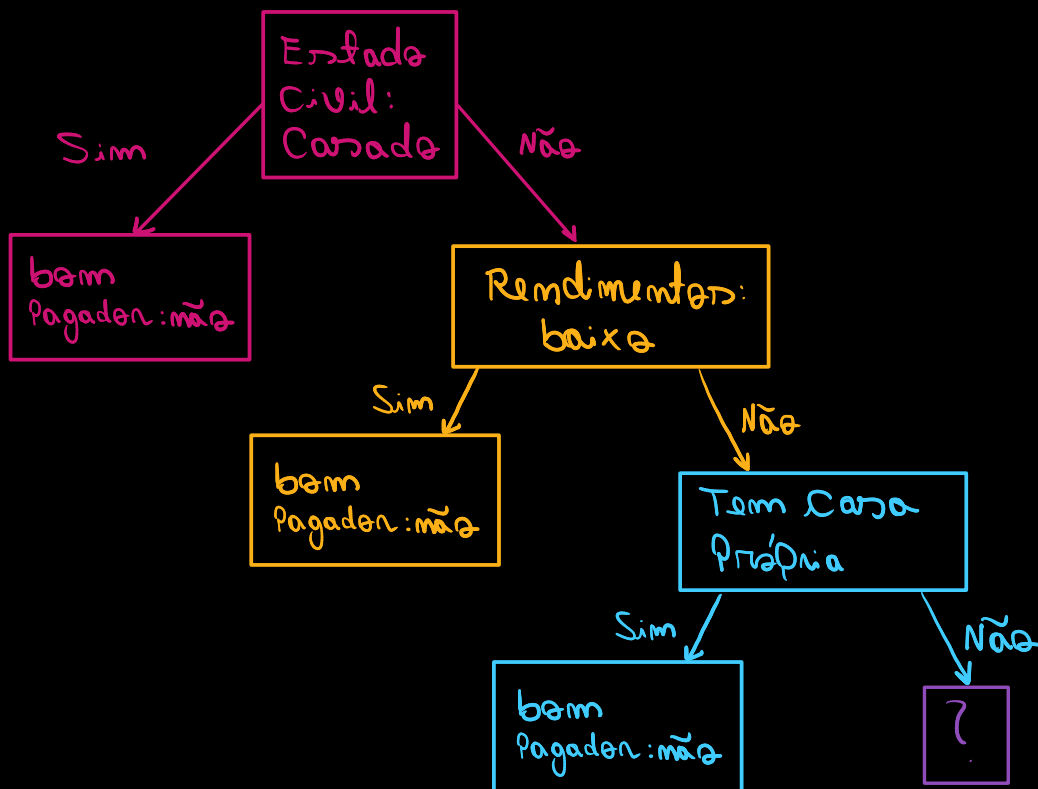
$$Gini(\text{Não Casa}) = 0$$

$$Gini(\text{Tem Casa Própria}) = \frac{3}{6} \left( 1 - \left( \frac{1}{3} \right)^2 - \left( \frac{2}{3} \right)^2 \right) + \frac{3}{6} \left( 1 - \left( \frac{3}{3} \right)^2 - \left( \frac{0}{3} \right)^2 \right) = 0,222$$

$$Gini(\text{Est. Civil: Solteiro | Divorciado}) = \frac{3}{6} \left( 1 - \left( \frac{2}{3} \right)^2 - \left( \frac{1}{3} \right)^2 \right) + \frac{3}{6} \left( 1 - \left( \frac{2}{3} \right)^2 - \left( \frac{1}{3} \right)^2 \right) = 0,444$$

$$Gini(\text{Rendimentos: Alto | Médio}) = \frac{2}{6} \left( 1 - \left( \frac{1}{2} \right)^2 - \left( \frac{1}{2} \right)^2 \right) + \frac{4}{6} \left( 1 - \left( \frac{3}{4} \right)^2 - \left( \frac{1}{4} \right)^2 \right) = 0,417$$

2ª Parte: O melhor índice Gini é o de tem casa própria: não.



## # 4ª iteração:

Registro	Tem casa própria	Estado Civil	Rendimentos	Bom Pagador
1	Sim	Solteiro	Alto	Não
2	Não	Casado	Médio	Não
3	Não	Solteiro	Baixo	Não
4	Sim	Casado	Alto	Não
5	Não	Divorciado	Médio	Sim
6	Não	Casado	Baixo	Não
7	Sim	Divorciado	Alto	Sim
8	Não	Solteiro	Médio	Sim
9	Não	Casado	Baixo	Não
10	Não	Solteiro	Médio	Sim
11	Sim	Divorciado	Médio	Não
12	Não	Divorciado	Alto	?

Desconsiderar amostras  
Com Tem Casa Própria: não.

## 1ª Parte:

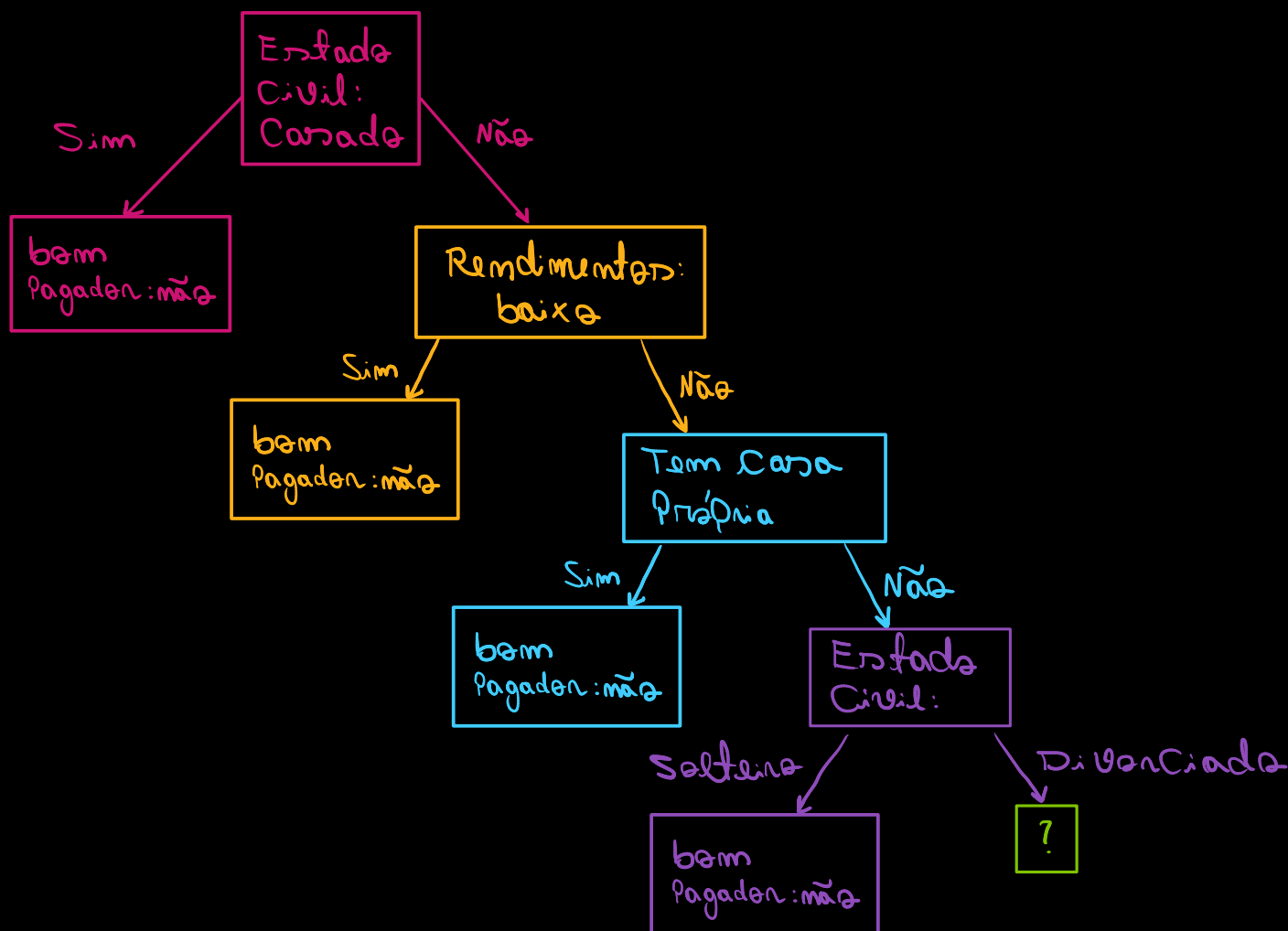
$$Gini(Tem Casa Própria) = \frac{3}{3} \left( 1 - \left( \frac{1}{3} \right)^2 - \left( \frac{2}{3} \right)^2 \right) = 0,444$$

$$Gini(Solteiro) = 0$$

$$Gini(Estado Civil: Solteiro | Divorciado) = \frac{1}{3} \left( 1 - \left( \frac{0}{1} \right)^2 - \left( \frac{1}{1} \right)^2 \right) + \frac{2}{3} \left( 1 - \left( \frac{1}{2} \right)^2 - \left( \frac{1}{2} \right)^2 \right) = 0,333$$

$$Gini(Rendimentos: Alto | Médio) = \frac{2}{3} \left( 1 - \left( \frac{1}{2} \right)^2 - \left( \frac{1}{2} \right)^2 \right) + \frac{1}{3} \left( 1 - \left( \frac{0}{1} \right)^2 - \left( \frac{1}{1} \right)^2 \right) = 0,333$$

2ª Parte: O mesmo índice Gini é o de Estado Civil: Solteiro



## #5ª iteração:

Registro	Tem casa própria	Estado Civil	Rendimentos	Bom Pagador
1	Sim	Solteiro	Alto	Não
2	Não	Casado	Médio	Não
3	Não	Solteiro	Baixo	Não
4	Sim	Casado	Alto	Não
5	Não	Divorciado	Médio	Sim
6	Não	Casado	Baixo	Não
7	Sim	Divorciado	Alto	Sim
8	Não	Solteiro	Médio	Sim
9	Não	Casado	Baixo	Não
10	Não	Solteiro	Médio	Sim
11	Sim	Divorciado	Médio	Não
12	Não	Divorciado	Alto	?

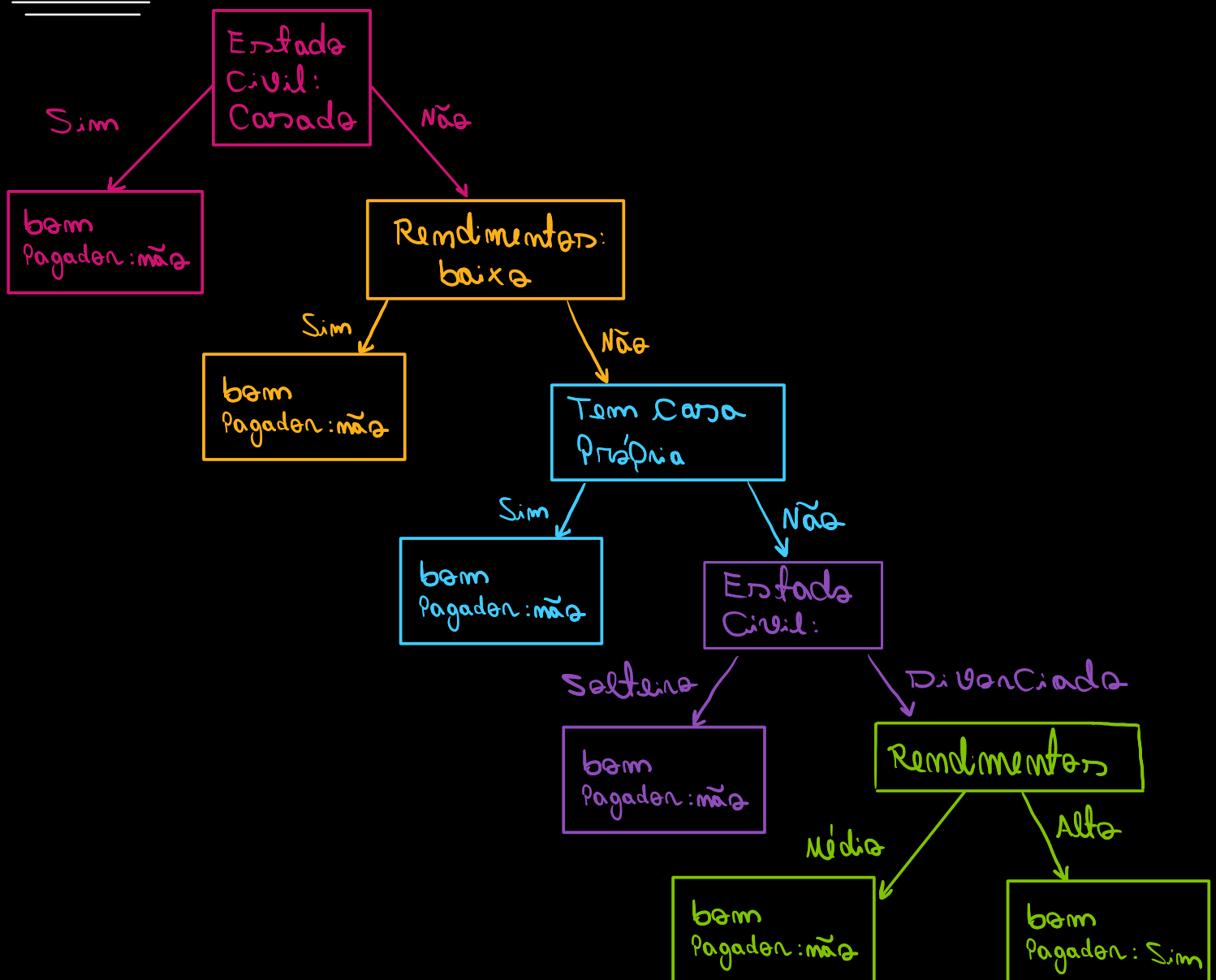
Desconsiderar amostras  
com Estado Civil: Solteiro

### 1ª Parada:

$$Gini(Tem Casa Própria) = \frac{2}{2} \left( 1 - \left( \frac{1}{2} \right)^2 - \left( \frac{1}{2} \right)^2 \right) = 0,5$$

$$Gini(Rendimentos: Alto | Médio) = \frac{1}{2} \left( 1 - \left( \frac{1}{1} \right)^2 - \left( \frac{0}{1} \right)^2 \right) + \frac{1}{2} \left( 1 - \left( \frac{0}{1} \right)^2 - \left( \frac{1}{1} \right)^2 \right) = 0$$

### 2ª Parada:



∴ A árvore decidiu que é um bom pagador