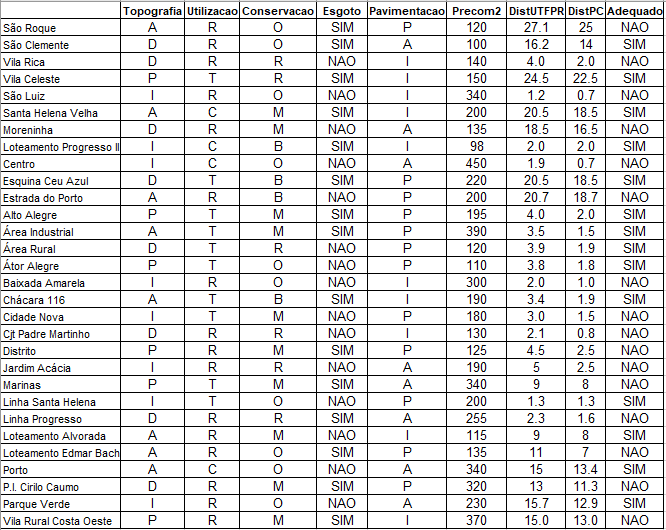
**Base de treino:**



**J48**

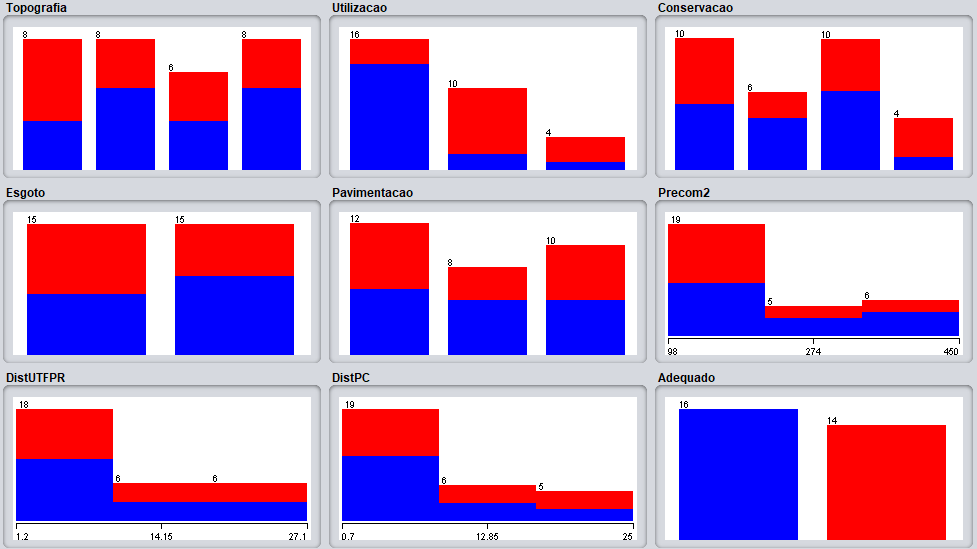
Algoritmo **J48**, testado em 7 casos de 10 instâncias cada. As instâncias são derivadas da base de treino (30 instâncias).

A taxa de acerto da base de treino é de 80% com 24 acertos, 6 erros e 0 não classificados.

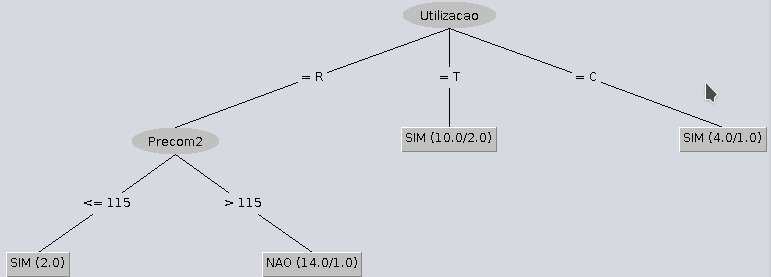
**Matriz de confusão da base:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 13 | 3 | A=NÃO |
| 3 | 11 | B=SIM |

Por 3 vezes o classificador inferiu que a resposta era NÃO e na verdade era SIM. E por 3 vezes o classificador inferiu que a resposta era SIM e na verdade era NÃO.

**Gráfico dos atributos:** 

**Árvore:**



Com base na árvore definida e no gráfico dos atributos, podemos analisar que o atributo **Precom2** mescla discretamente as saídas de **SIM** e **NÃO**, permitindo uma faixa de erro.

**Tabela de testes:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| TESTE | ACERTOS | ERROS | %ACERTOS | %SIM | %NÃO |
| 1 | 9 | 1 | 90 | 83,3 | 100 |
| 2 | 9 | 1 | 90 | 83,3 | 100 |
| 3 | 9 | 1 | 90 | 83,3 | 100 |
| 4 | 9 | 1 | 90 | 83,3 | 100 |
| 5 | 7 | 3 | 70 | 62,5 | 100 |
| 6 | 7 | 3 | 70 | 50 | 100 |
| 7 | 9 | 1 | 90 | 83,3 | 100 |
| **TOTAIS** | **8,428571** | **1,571429** | **84,28571** | **75,57143** | **100** |

**Modificações:**

**Teste 01:**

**Teste 02:**

Linha 4 - Precom2 -> de 140 para 290

Linha 8 - Precom2 -> de 135 para 200

**Teste 03:**

Linha 5 - DistUTFPR -> de 150 para 30

Linha 9 - DistUTFPR -> de 98 para 290

**Teste 04:**

Linha 7 - DistPC -> de 18.5 para 14.3

Linha 8 - DistPC -> de 16.5 para 3.0

**Teste 05:**

Linha 2 - Precom (de 120 para 80) - DistUTFPR (de 27.1 para 18.0) - DistPC (de 25 para 10)

Linha 4 - Precom (de 140 para 100) - DistUTFPR ( de 4.0 para 12.5) - DistPC (de 2.0 para 9.0)

Linha 5 - Precom (de 150 para 100)

Linha 8 - DistUTFPR (de 18.5 para 9.6)

Linha 9 - DistPC de 2.0 para 25.0)

**Teste 06:**

Linha 4 - Conservacao ( de R para O) - Pavimentacao ( I para A )

Linha 7 - Conservacao ( de M para O) - Adequado ( de SIM para NAO )

Linha 11 - Topografia (de D para A) - Adequado (de SIM para NAO) - Pavimentacao ( de P para A )

**Teste 07:**

Linha 5 - Utilizacao (de T para C) - Precom2 (de 150 para 200)

Linha 6 - Utilizacao (de R para T) - Precom2 (de 340 para 120)

Linha 7 - Utilizacao (de C para R) - Precom2 (de 200 para 360)

**Matriz de confusão dos testes:**

**Teste 01:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 4 | 1 | A=NÃO |
| 0 | 5 | B=SIM |

**Teste 02:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 4 | 1 | A=NÃO |
| 0 | 5 | B=SIM |

**Teste 03:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 4 | 1 | A=NÃO |
| 0 | 5 | B=SIM |

**Teste 04:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 4 | 1 | A=NÃO |
| 0 | 5 | B=SIM |

**Teste 05:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 2 | 3 | A=NÃO |
| 0 | 5 | B=SIM |

**Teste 06:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 4 | 3 | A=NÃO |
| 0 | 3 | B=SIM |

**Teste 07:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 4 | 1 | A=NÃO |
| 0 | 5 | B=SIM |

Para os casos de teste acima mostrados, o algoritmo J48 teve uma taxa de 84,2% de acertos. Nos casos em que a resposta foi **NÃO**, o algoritmo acertou 100%.

**ID3**

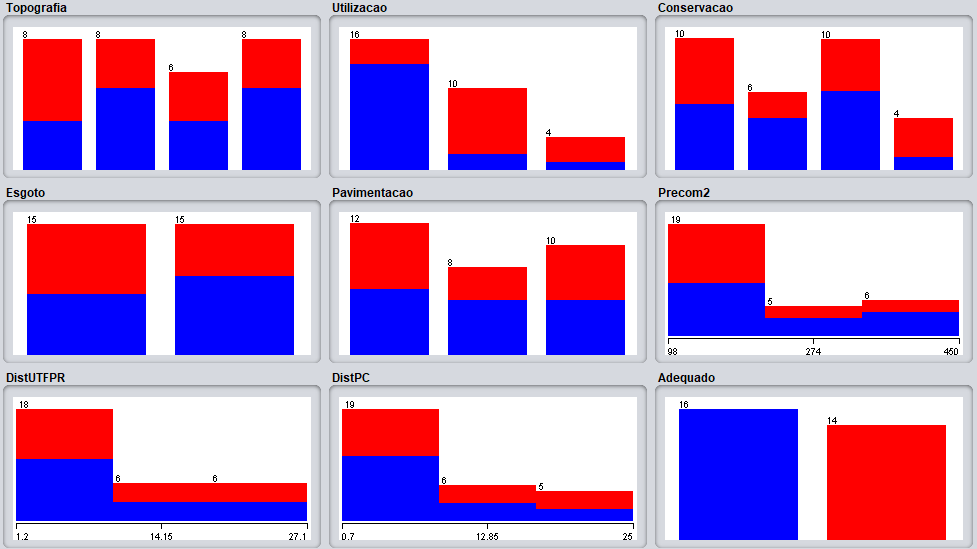
Algoritmo **ID3**, testado em 3 casos de 10 instâncias cada. As instâncias são derivadas da base de treino (30 instâncias).

A taxa de acerto da base de treino é de 63,3% com 19 acertos, 8 erros e 3 não classificados.

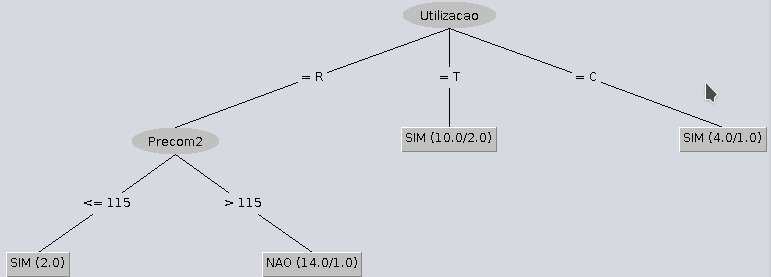
**Matriz de confusão da base:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 11 | 3 | A=NÃO |
| 5 | 8 | B=SIM |

Por 3 vezes o classificador inferiu que a resposta era SIM e na verdade era NÃO. E por 5 vezes o classificador inferiu que a resposta era NÃO e na verdade era SIM.

**Gráfico dos atributos:** \*\*\*\*\*\*\*\*\*\*

**Árvore:**



**Tabela de testes:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| TESTE | ACERTOS | ERROS | %ACERTOS | %SIM | %NÃO |
| 1 | 10 | 0 | 100 | 100 | 100 |
| 2 | 10 | 0 | 100 | 100 | 100 |
| 3 | 10 | 0 | 100 | 100 | 100 |
| **TOTAIS** | **10** | **0** | **100** | **100** | **100** |

**Modificações:**

**Teste 01:**

Linha 3 - Topografia (de D para A)

Linha 6 - Topografia (I para D)

Linha 11 - Topografia (D para I)

**Teste 02:**

Linha 4 - Conservacao (de R para O)

Linha 6 - Conservacao (de O para M)

Linha 9 - Conservacao (de B para M)

**Teste 03:**

Linha 5 - Esgoto (de SIM para NAO) - Pavimentacao (de I para P)

Linha 9 - Esgosto (de SIM para NAO) - Pavimentacao (de I para A)

Linha 10 - Esgosto (de NAO para SIM) - Pavimentacao (de A para I)

**Matriz de confusão dos testes:**

**Teste 01:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 5 | 0 | A=NÃO |
| 0 | 5 | B=SIM |

**Teste 02:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 5 | 0 | A=NÃO |
| 0 | 5 | B=SIM |

**Teste 03:**

|  |  |  |
| --- | --- | --- |
| A | B |  |
| 5 | 0 | A=NÃO |
| 0 | 5 | B=SIM |

Para os casos de teste acima mostrados, o algoritmo ID3 teve uma taxa de 100% de acertos.