**Aula Prática 5**

**Tema: Árvore de Decisão**

**Murielly Oliveira Nascimento – 11921BSI222**

Perguntas:

1. Considerando o algoritmo Árvore de Decisão e a base de dados mostrada a seguir, classifique o dado exemplo de teste. Use a medida de erro de classificação para construir a árvore.

Mostre o passo a passo da execução do seu algoritmo.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Atributos | Instâncias | Frequência | Yes | No | Erro de Classificação | Média Ponderada | Ganho |
| Outlook | Sunny | 5 | 2 | 3 | 0,4 | 0,285714286 | 0,714286 |
| Overcast | 4 | 4 | 0 | 0 |  |  |
| Rainy | 5 | 3 | 2 | 0,4 |  |  |
| Temp | Hot | 4 | 2 | 2 | 0,5 | 0,357142857 | 0,642857 |
| Mild | 6 | 4 | 2 | 0,333333333 |  |  |
| Cool | 4 | 3 | 1 | 0,25 |  |  |
| Humidity | High | 7 | 3 | 4 | 0,428571429 | 0,285714286 | 0,714286 |
| Normal | 7 | 6 | 1 | 0,142857143 |  |  |
| Windy | False | 8 | 6 | 2 | 0,25 | 0,357142857 | 0,642857 |
| True | 6 | 3 | 3 | 0,5 |  |  |
| Play | No | 5 |  |  | 0,642857143 | 1 |  |
| Yes | 9 |  |  | 0,357142857 |  |  |

1. Procure uma base no site do UCI que contenha atributos nominais e numéricos. Execute usando o Weka o algoritmo de árvore de decisão na base escolhida. Plote a árvore. Converta os atributos numéricos para categóricos e execute novamente o algoritmo. Plote a nova árvore. Compare os resultados. O que você conseguiu identificar de diferente nas árvores? Lembre-se, se a base possuir valores ausentes, você deverá pré-processá-la inicialmente. Procure por bases não muito grandes a fim de facilitar a visualização da árvore.

A base usada para as análises foi a íris. A árvore de decisão gerada a partir dela é como segue:

Diagrama

Descrição gerada automaticamente

Convertendo os atributos numéricos para categóricos obtemos a seguinte árvore:

Diagrama

Descrição gerada automaticamente

Na discretização os atributos numéricos são divididos em n intervalos, nesse caso foram três. Como na árvore de decisão o objetivo é que em cada folha todos os dados sejam da mesma classe, nota-se que para esse caso o uso desse pré-processamento auxiliou a execução do algoritmo de classificação.