**INSTITUTO POLITÉCNCO NACIONAL**

**ESCUELA SUPERIOR DE CÓMPUTO**

**MATERIAL EDUCATIVO PARA LA UNIDAD DE APRENDIZAJE DE MINERÍA DE DATOS. 2022–2**

*Grupo 5CDM1*

**PRACTICA DE ÁRBOLES DE DECISIÓN**

|  |
| --- |
| Nombres: |
|  |
|  |

1. **Descripción del conjunto de datos.**

Autores del conjunto de datos:

Donante: Ronny Kohavi and Barry Becker. Data Mining and Visualization. Silicon Graphics. e-mail: ronnyk '@' live.com for questions.

Enlace de acceso: <https://archive.ics.uci.edu/ml/datasets/adult>

1. **Objetivo de la práctica.**

Realizar un árbol de decisión para predecir si una persona gana más de 50K al año.

1. **Diccionario de datos.**

Construya el diccionario de datos considerando la siguiente estructura.

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Nombre** | **Tipo** | **Dominio** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

1. **Resultados**

Presente los resultados considerando lo siguiente:

* Realice y describa los resultados de cinco consultas descriptivas en el conjunto de datos
* Presente propiedades estadísticas del conjunto de datos
* Describa las medidas generadas a partir de la matriz de confusión (archivo anexo)
* Analice este comportamiento en función la cantidad de elementos de cada tipo que existen en el conjunto de datos
* Anexe el modelo y las reglas generadas.

|  |  |
| --- | --- |
| Use los siguientes nodos: |  |