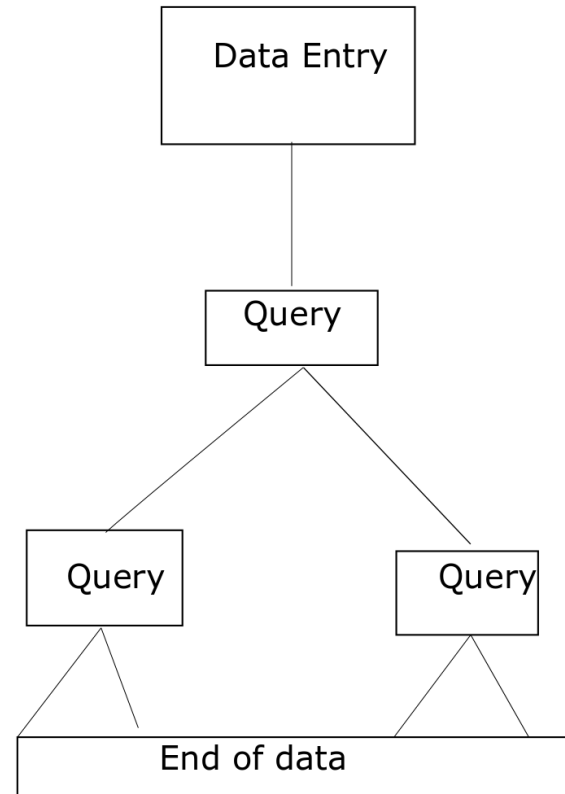


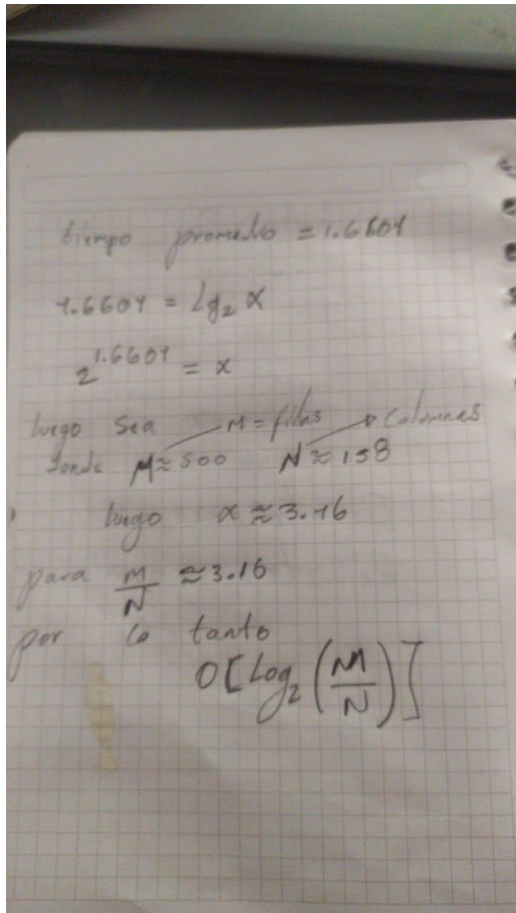
The Success of the students: A study based on test results

Andrés Felipe Téllez Rodríguez
Jonathan Andrés Bedoya Zapata

Estructura de datos



Complejidades



tiempo promedio = 1.6604

$$1.6604 = \log_2 X$$
$$2^{1.6604} = X$$

luego sea $M = \text{filas}$ y $N = \text{columnas}$
donde $M \approx 500$ y $N \approx 158$

luego $X \approx 3.16$

para $\frac{M}{N} \approx 3.16$

por lo tanto

$$O\left[\log_2\left(\frac{M}{N}\right)\right]$$

$$O(\log(M/N))$$

Donde M son Filas Y N son columnas

Código en funcionamiento

```
main.py
1 from __future__ import print_function
2 import time
3 import pandas as pd
4
5 # Este método sirve para cargar los datos de prueba y
  training_data para el árbol de decisión
6 # además permite obtener el header del csv para utilizarlas
  en las preguntas
7 def carga_datos(ruta_entrenamiento,ruta_prueba):
8     datos_entrenamiento = []
9     datos_prueba = []
10    Filas = []
11
12    #Cargar el archivo de datos
13    archivo_entrenamiento = open(ruta_entrenamiento,
14                                 encoding='utf-8')
15    archivo_prueba = open(ruta_prueba,encoding='utf-8')
16
17    #Crear una lista de listas de estudiantes
18    for linea in archivo_entrenamiento:
19        linea = linea[:-1] #Se elimina el salto de línea en
20        el csv
21        Filas = linea.split(";")
22        datos_entrenamiento.append(Filas)
23
24    #Retirar las cabeceras del data_array, estas se utilizan
25    para imprimir el árbol
26    cabeceras = datos_entrenamiento[0]
```

```
Is cole_cod_dane_sede == 47?
--> True:
    Predict {'0': 2}
--> False:
    Is cole_area_ubicacion == 4?
    --> True:
        Predict {'0': 1}
    --> False:
        Is cole_caracter == 0?
        --> True:
            Predict {'0': 1}
        --> False:
            Predict {'1': 11}
--> False:
    Is periodo.1 == Técnica o tecnológica completa?
    --> True:
        Predict {'1': 3}
    --> False:
        Is cole_genero == 54?
        --> True:
            Predict {'1': 2}
        --> False:
            Is cole_area_ubicacion == 55?
            --> True:
                Predict {'1': 1}
            --> False:
                Is cole_nombre_sede == B1?
                --> True:
                    Predict {'1': 1}
                --> False:
                    Predict {'0': 28}
Tiempo promedio de ejecución: 1.0005805492401123
Exactitud: 75.0%
```

Gracias por su atención