



PREDICCIÓN DEL ÉXITO EN LAS PRUEBAS SABER 11

Presentación del Equipo



Katherin Nathalia
Allin Murillo



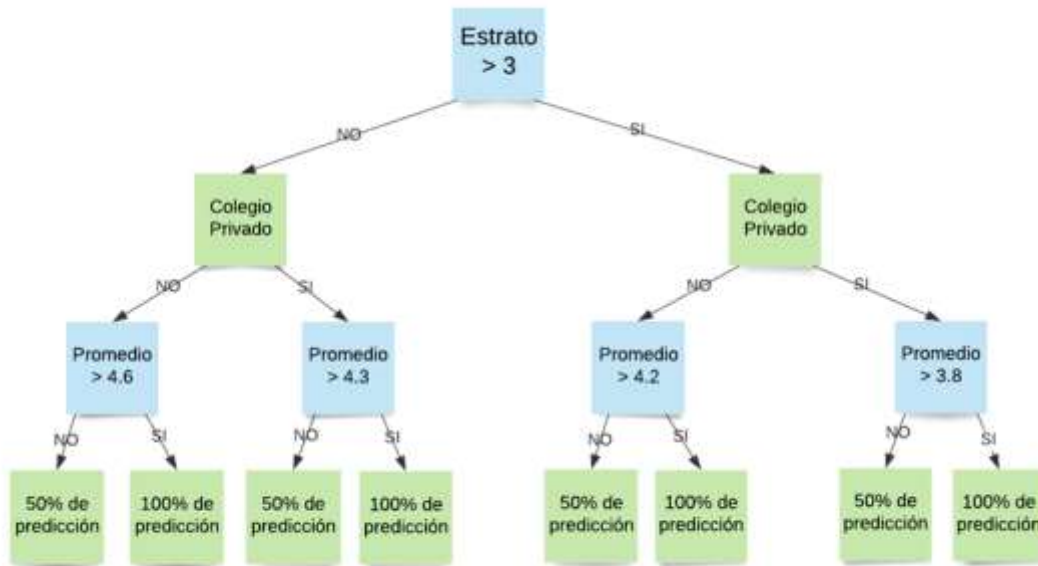
Alberto Andrés
Díaz Mejía



<https://github.com/AlbertoD10-edu/ST0245-002/tree/master/proyecto>

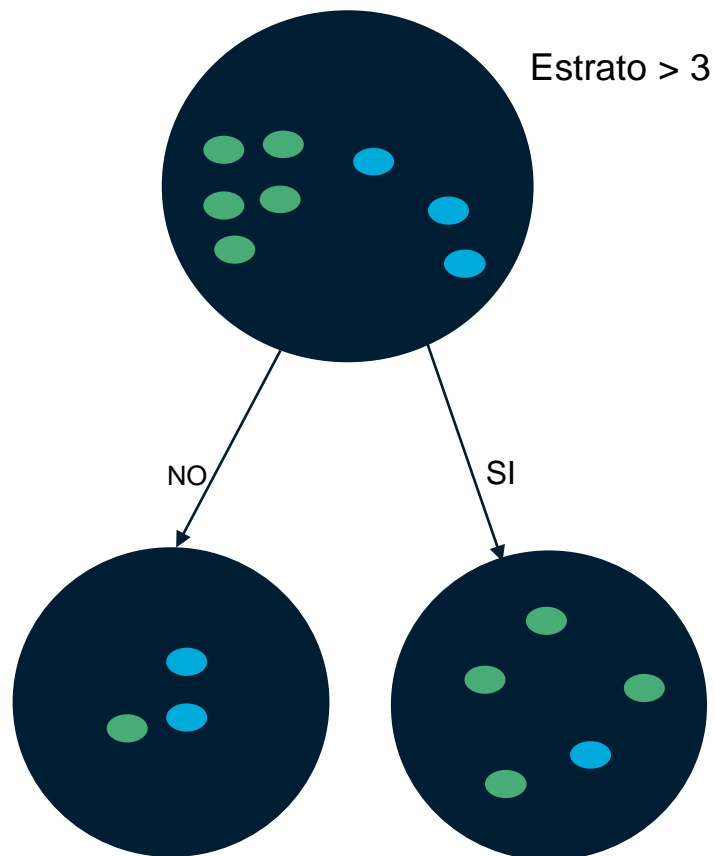


Diseño del Algoritmo



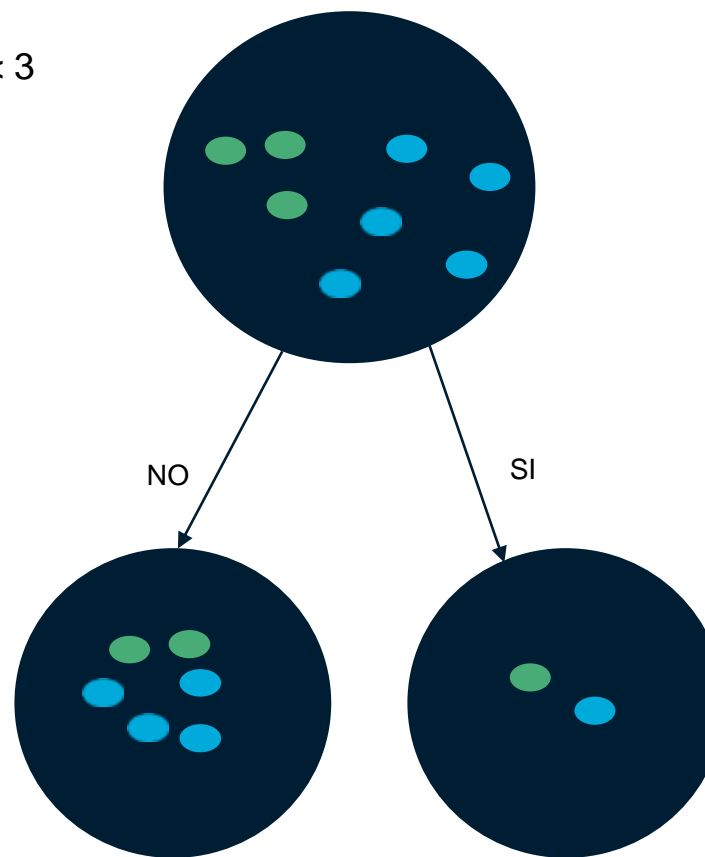
El tipo de algoritmo que decidimos implementar es el CART, nos basamos en la influencia que tiene el tipo de estrato y la educación a la que tienen acceso los alumnos, ya que estos son factores muy importantes para predecir el éxito de las pruebas; claro está que hay muchos factores que se pueden usar para una mejor predicción como promedio por materia, etc.

División de un nodo



Esta gráfica está basada en “Estrato > 3 ”. Para este caso la impureza Gini de la izquierda es 0,16, la impureza Gini de la derecha es 0,08 y la impureza ponderada es 0,12.

Estrato ≤ 3



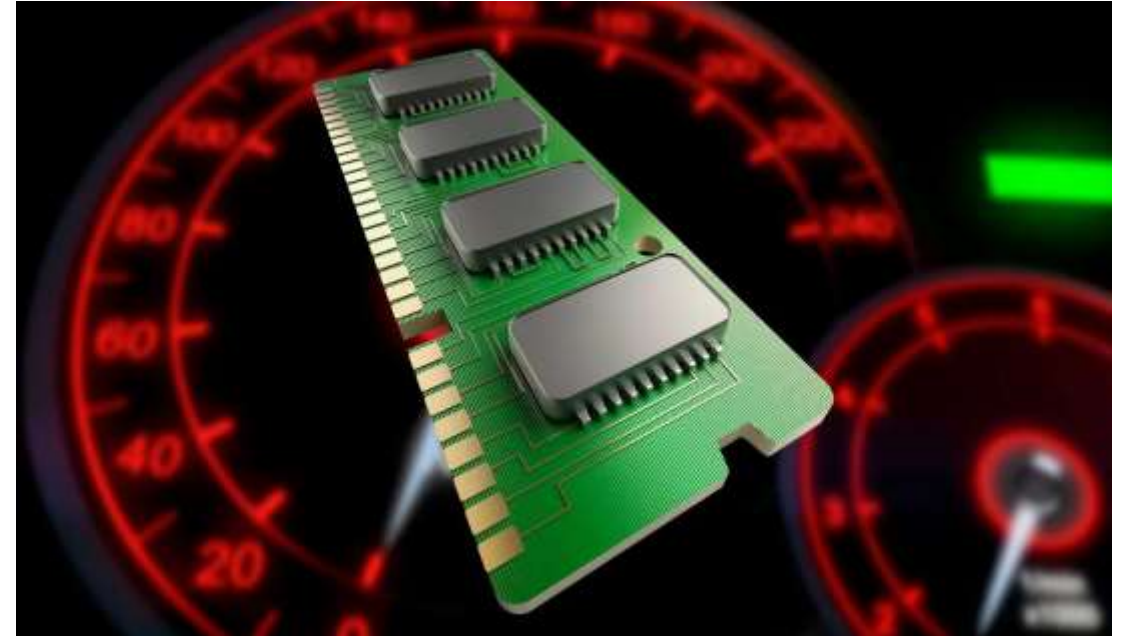
Esta gráfica está basada en “Estrato ≤ 3 ”. Para este caso la impureza Gini de la izquierda es 0,32, la impureza Gini de la derecha es 0,08 y la impureza ponderada es 0,20.

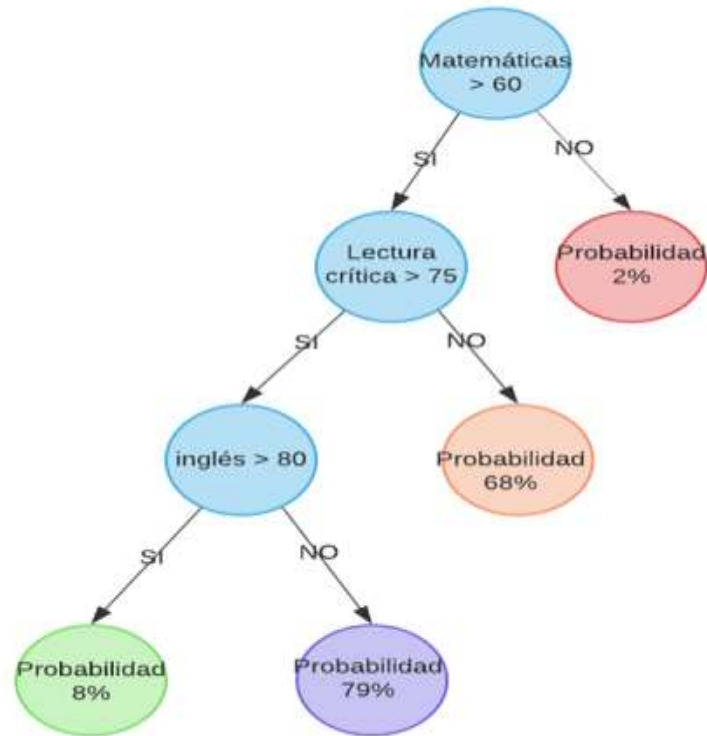
Complejidad del Algoritmo



-----	Complejidad en tiempo	Complejidad en memoria
Algoritmo en práctica	$O(\log n)$	$O(n^2)$

Complejidad en tiempo y memoria del algoritmo CART, notamos que la complejidad en memoria necesita realizar una iteración por todos los elementos en cada uno de los elementos a procesar y la complejidad en tiempo vimos que es como si "troceara" el problema para abordarlo.





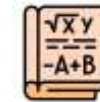
Características Más Relevantes



Lectura Crítica



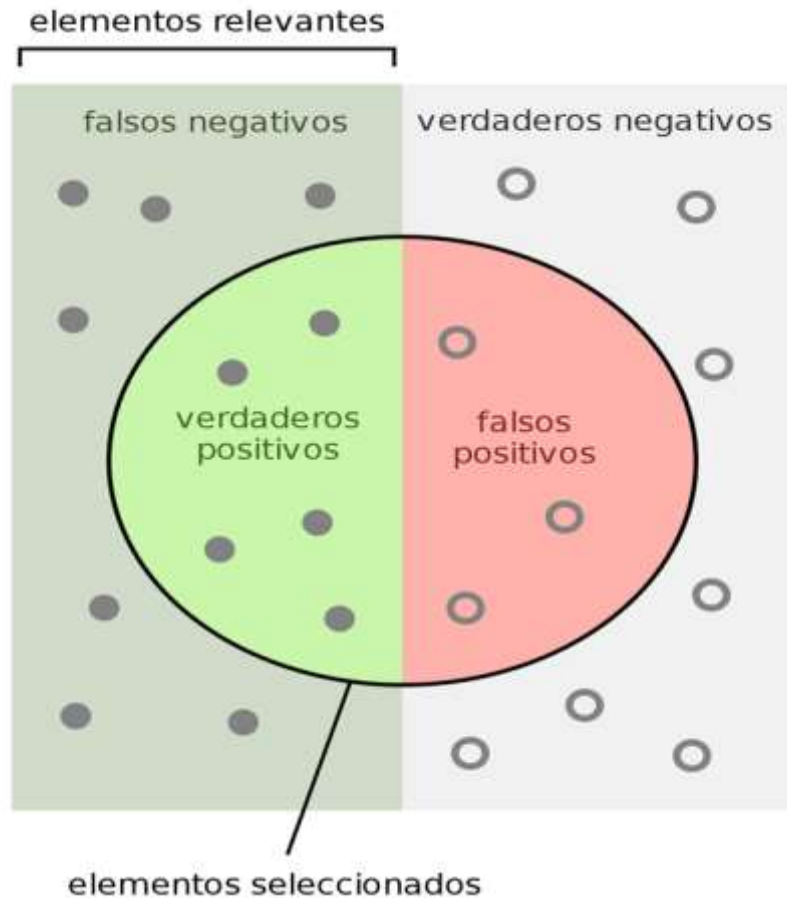
Inglés



Matemáticas

Nos guiamos con un estudio encontrado en el siguiente enlace http://rdigitales.uptc.edu.co/memorias/index.php/apli_estad/apli_esta2016/paper/viewFile/2169/2232, este estudio fue hecho hace un par de años y decidimos escoger estas 3 materias en específico porque son las que presentan mas bajos porcentajes en los niveles altos

Métricas de Evaluación



$$\text{Sensibilidad} = \frac{\text{Número de alumnos identificados}}{\text{Número de alumnos registrados}}$$

$$\text{Precisión} = \frac{\text{Número de alumnos con éxito}}{\text{Número de alumnos con predicción de éxito}}$$

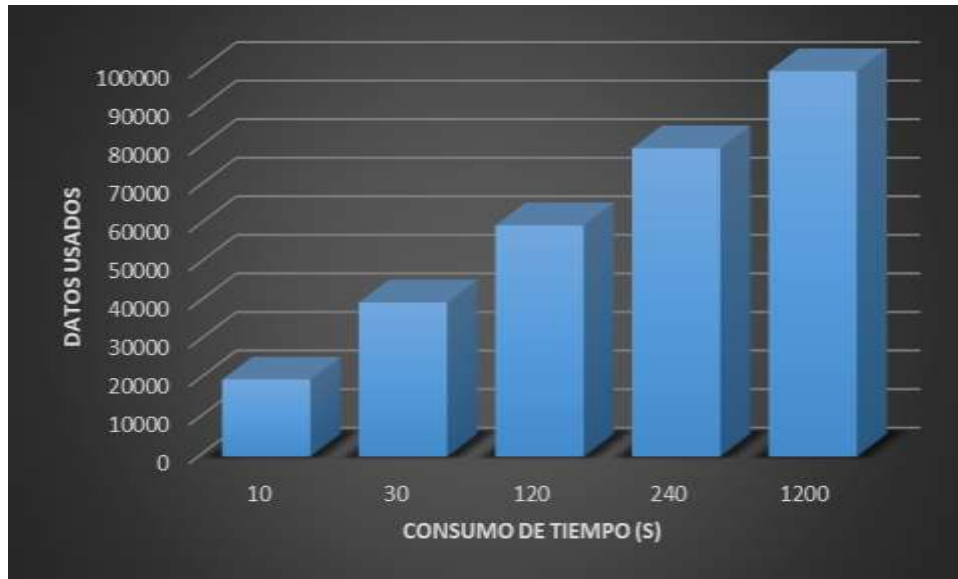
$$\text{Exactitud} = \frac{\text{Número de entradas}}{\text{Predicciones correctas}}$$

-----	Conjunto de entrenamiento	Conjunto de validación
Exactitud	0,7	0,79
Precisión	0,6	0,8
Sensibilidad	0,55	0,38

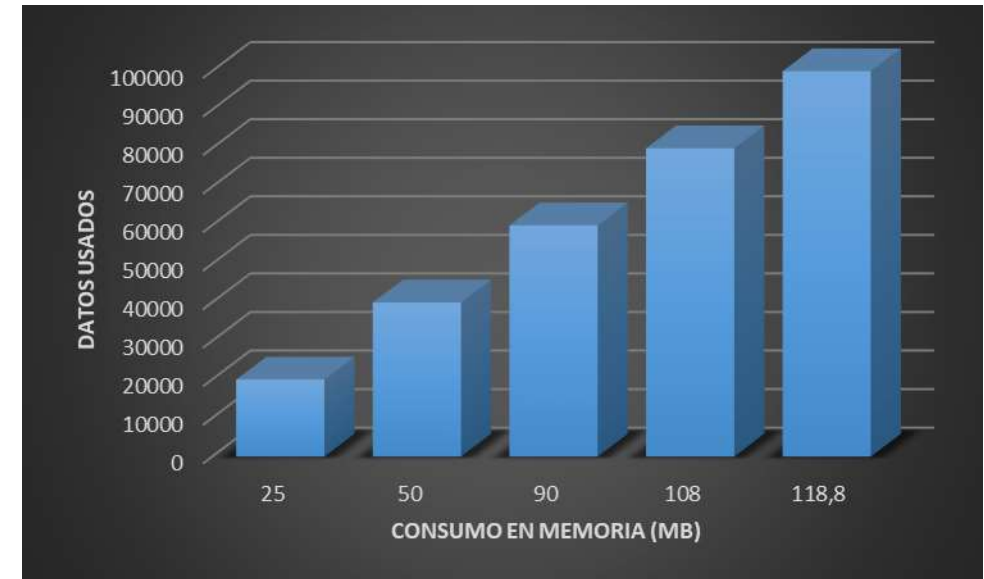


Métricas de evaluación obtenidas con el conjunto de datos de entrenamiento de 100,000 estudiantes y el conjunto de datos de validación de 2,000 estudiantes.

Consumo de tiempo y memoria



 Consumo de tiempo



 Consumo de memoria



¡GRACIAS!