A man with a beard is sitting at a desk in a library, reading an open book. He is wearing a light blue sweater. The background shows bookshelves filled with books. The lighting is soft and focused on the man and his book.

LA PREDICCIÓN DEL ÉXITO ACADÉMICO A TRAVÉS DE LAS PRUEBAS SABER

Presentación del Equipo



Juan P.
Henao



Diego
Vanegas



Miguel
Correa

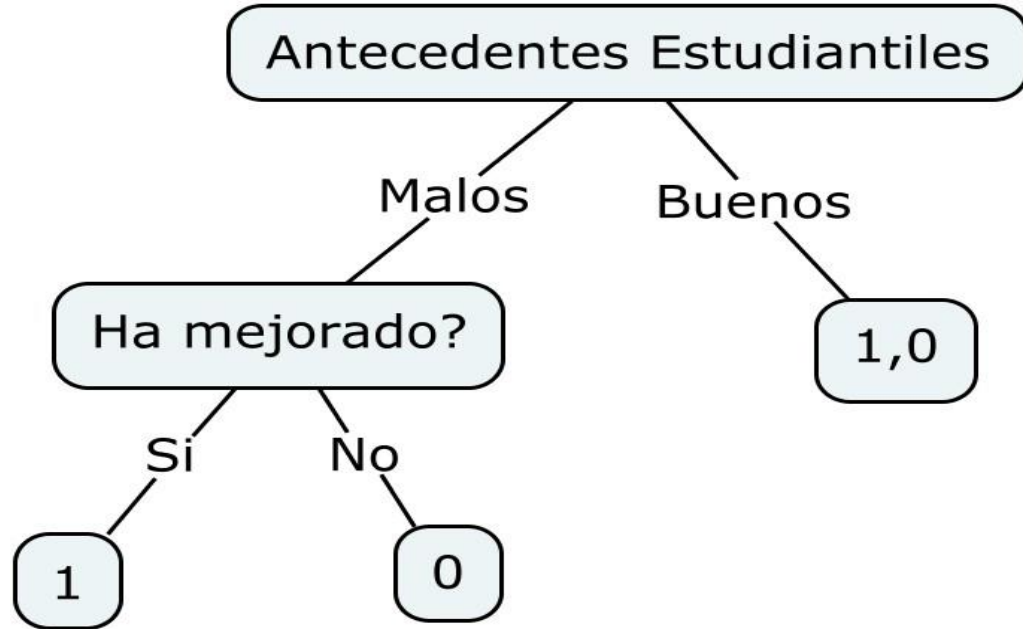


Mauricio
Toro

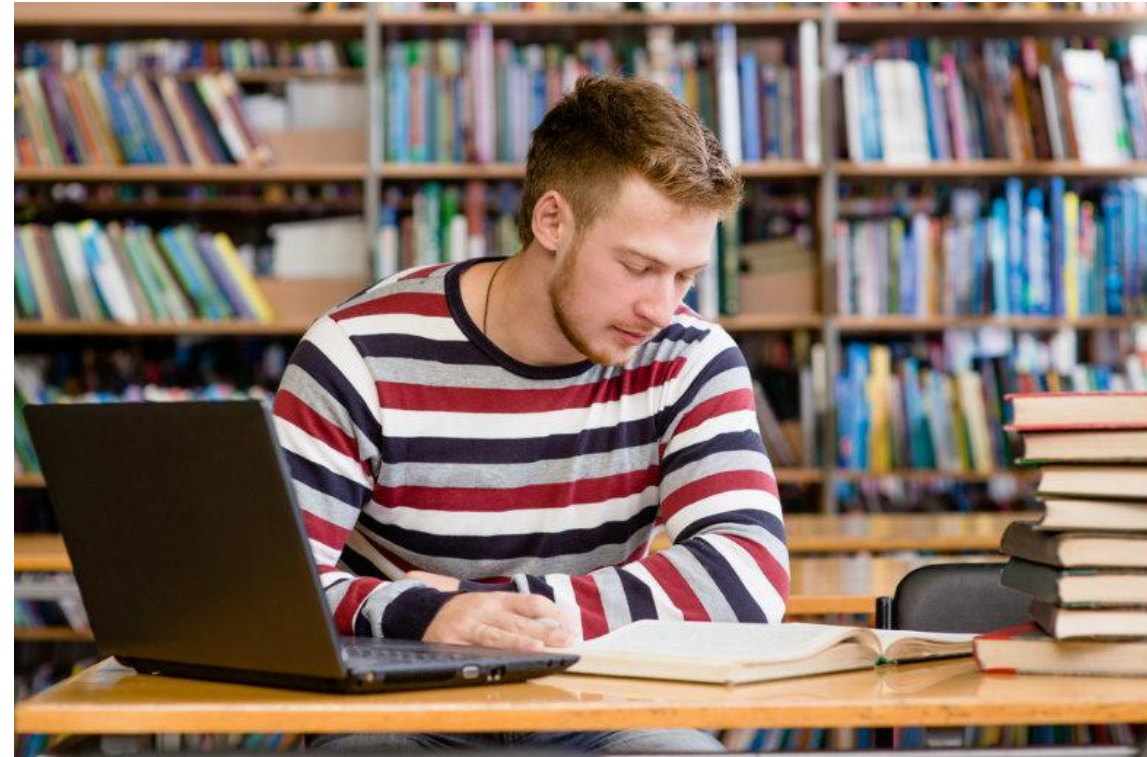


<http://github.com/Juanp0128/ST0245-001/tree/master/proyecto>





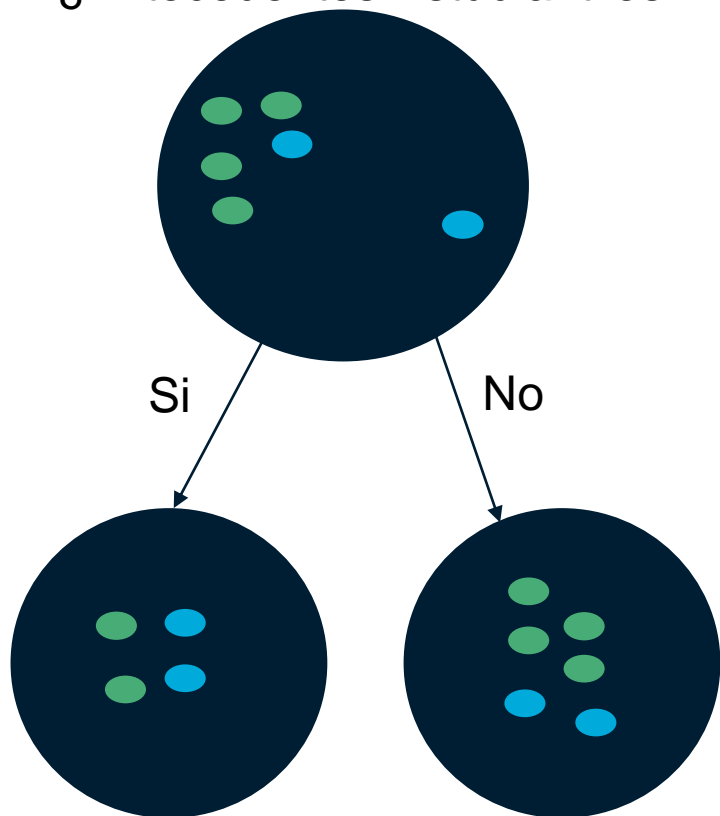
Podemos reducir la impureza mediante los antecedentes estudiantiles de forma considerable, donde sabemos que si un estudiante tiene antecedentes negativos, se verá reflejado un posible bajo promedio en los resultados. Mientras que, si el estudiante ha tenido buenos antecedentes, es más probable que el resultado sea encima del promedio, aunque no podríamos asegurar con certeza, ya que no sabemos cómo fueron los métodos evaluativos tanto del Colegio como de la Universidad, y no sabemos si el estudiante fue honesto en sus procesos educativos.



División de un nodo

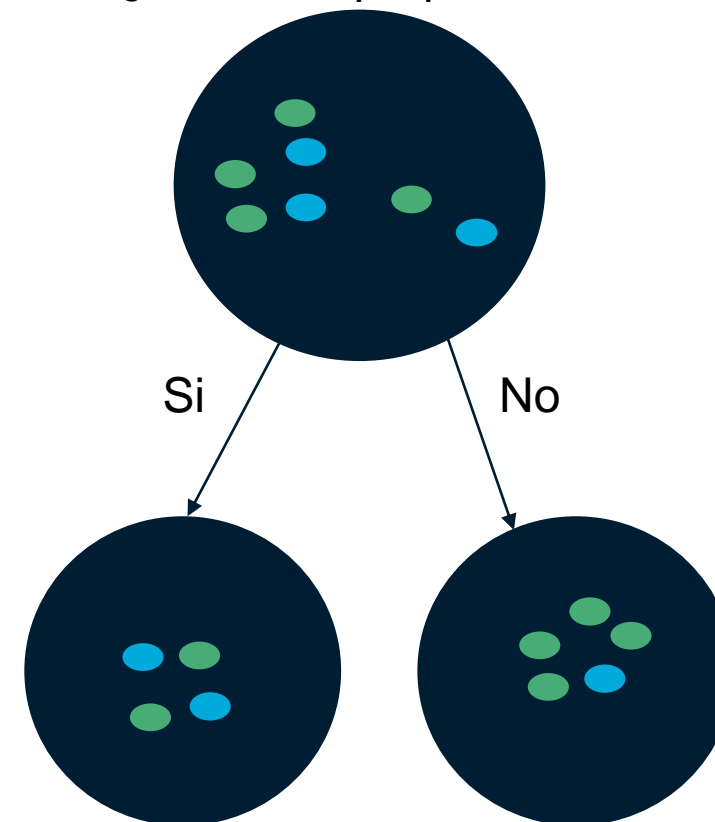


¿Antecedentes Estudiantiles?



Esta división está basada en la condición "Antecedentes Estudiantiles" Para este caso, la impureza Gini de la izquierda es 0.48, la impureza Gini de la derecha es 0.50 y la impureza ponderada es de 0.98.

¿Curso de preparación?



Esta división está basada en la condición "Curso Preparacion" Para este caso, la impureza Gini de la izquierda es 0.49, la impureza Gini de la derecha es 0.50 y la impureza ponderada es 0.99.

Complejidad del Algoritmo

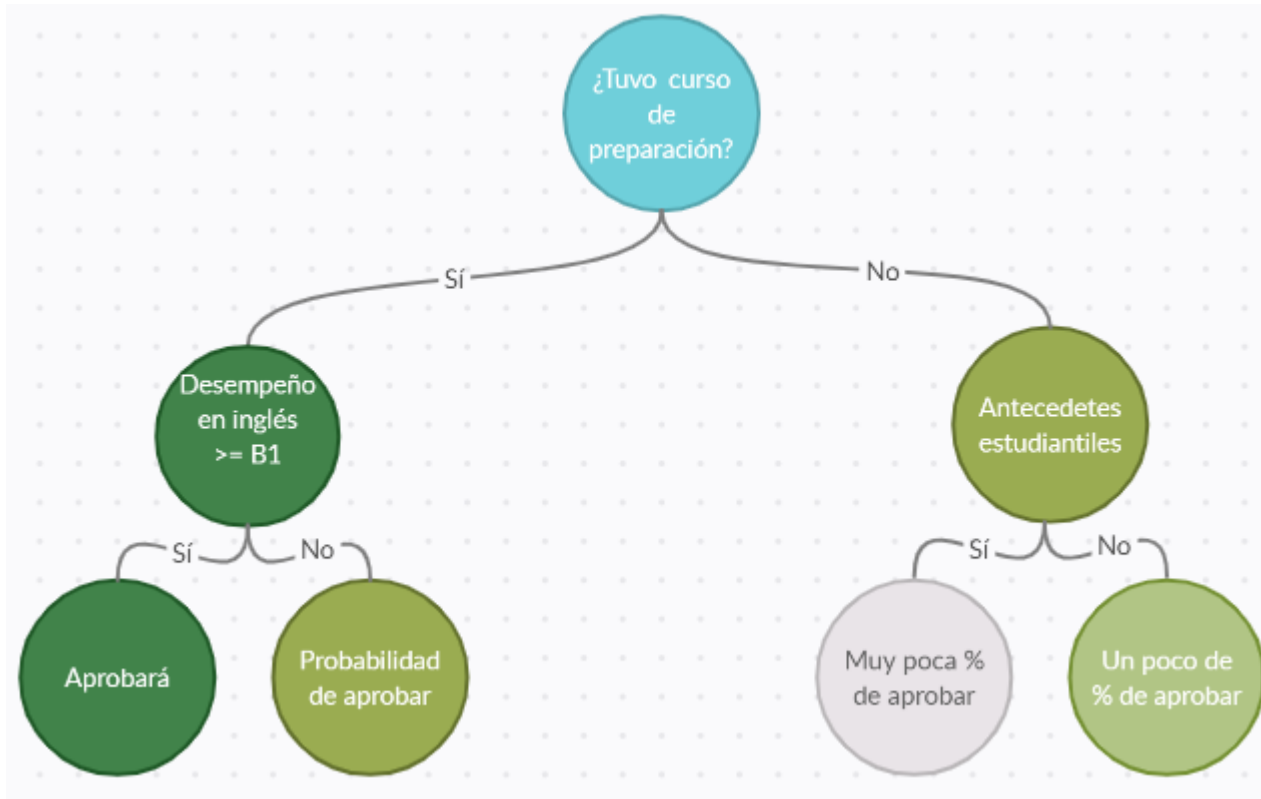


	Complejidad en tiempo	Complejidad en memoria
Entrenamiento del modelo	$O(N \cdot \log(M))$	$O(N \cdot M)$
Validación del modelo	$O(N \cdot M)$	$O(N \cdot M)$

N como la cantidad de estudiantes, y M como la cantidad de columnas donde es un valor exponencial



Modelo de Árbol de Decisión



Un árbol de decisión para predecir el resultado del Saber Pro usando los resultados del Saber 11. El azul representa el nodo padre (Pregunta principal) y el verde, dependiendo de su tonalidad si tiende a ser más oscuro, mayor probabilidad de éxito tiene el estudiante.

Características Más Relevantes



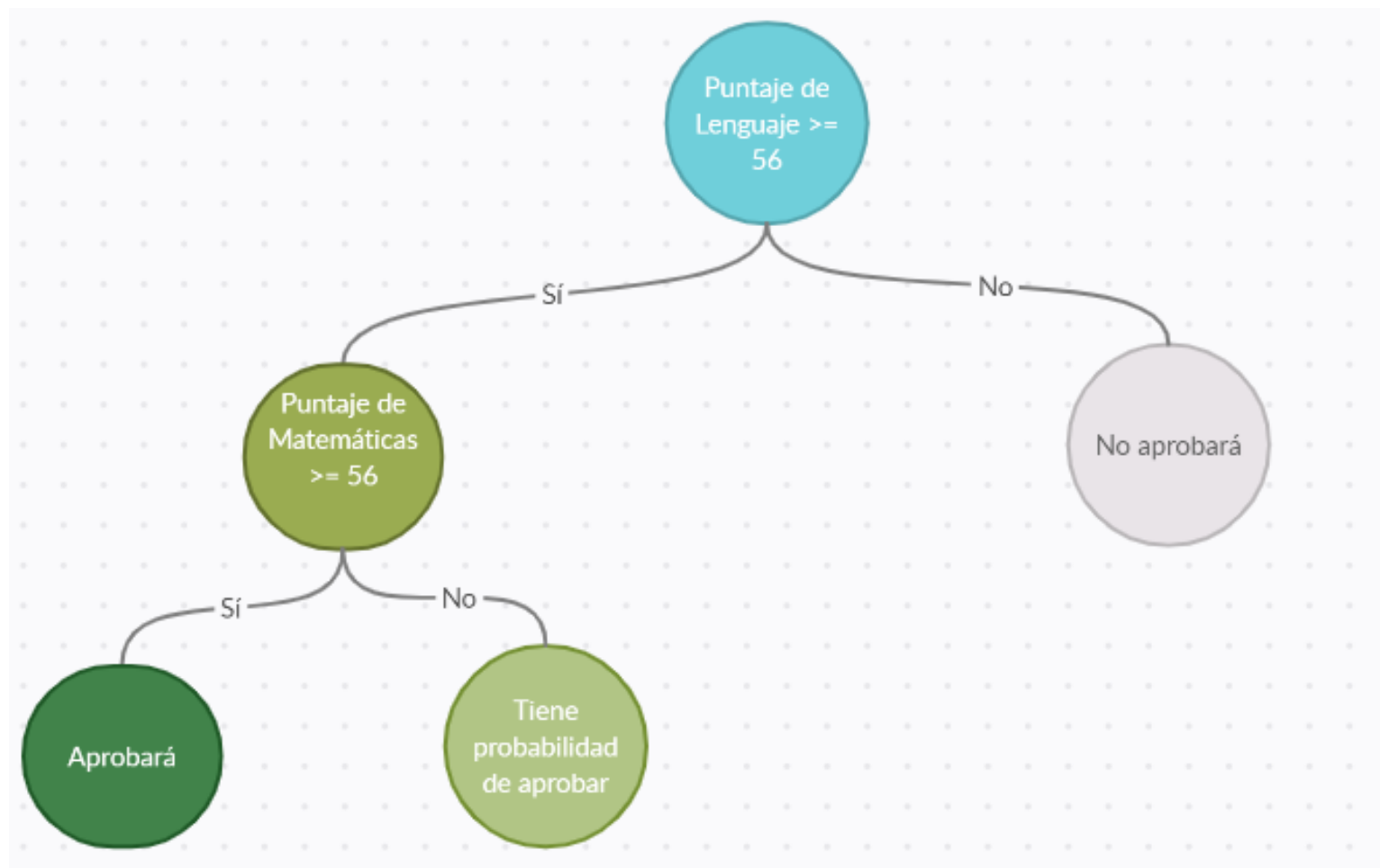
Lenguaje y Matemáticas



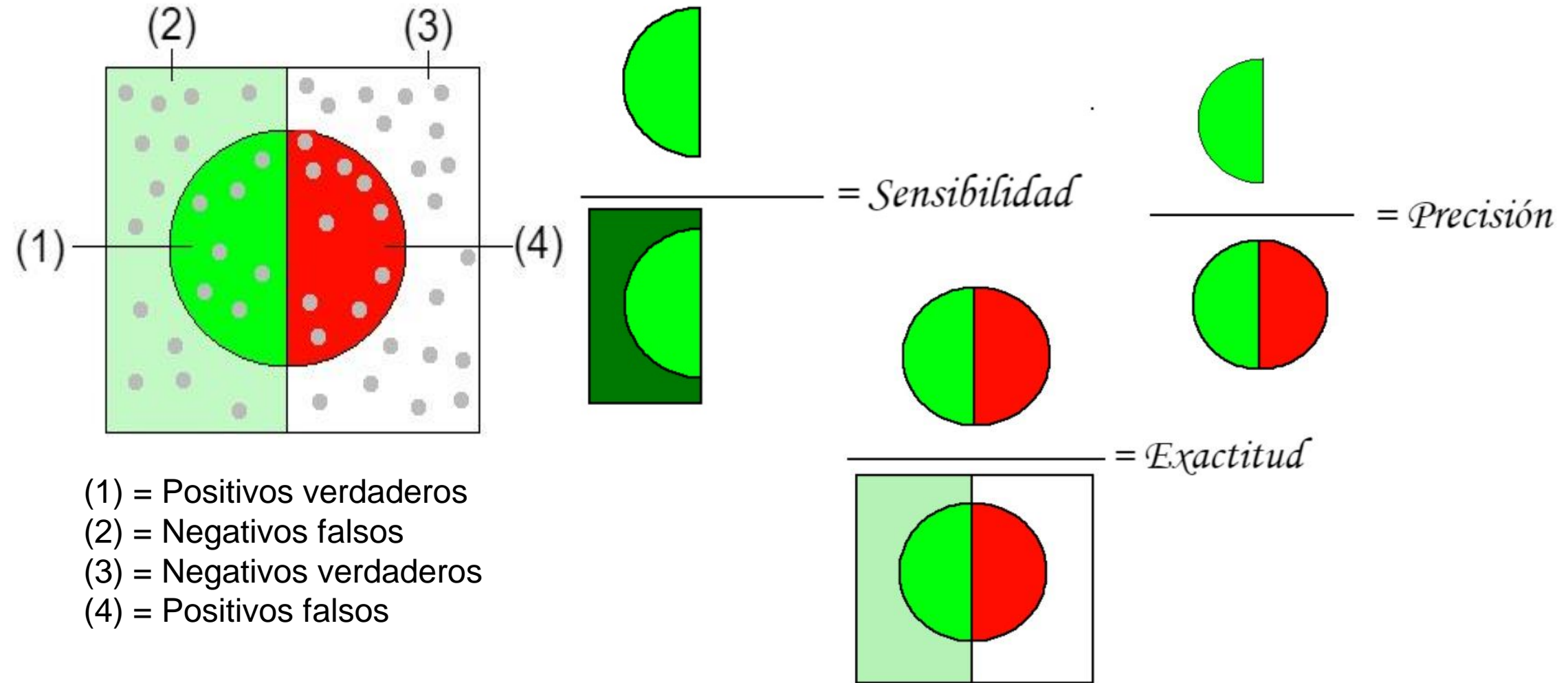
Inglés



Antecedentes



Métricas de Evaluación



- (1) = Positivos verdaderos
- (2) = Negativos falsos
- (3) = Negativos verdaderos
- (4) = Positivos falsos

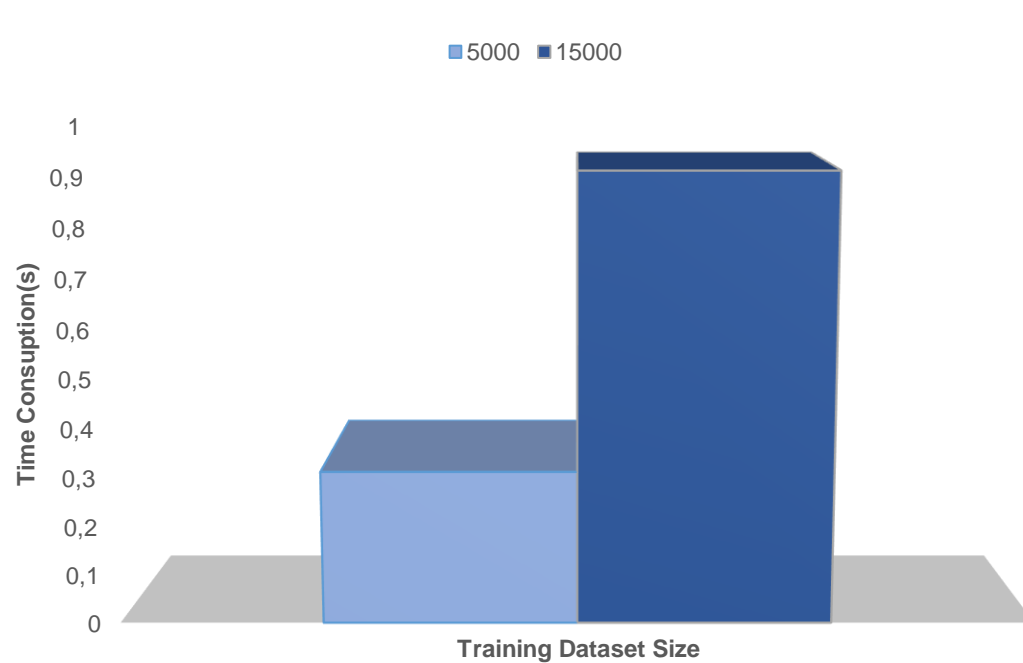


	Conjunto de entrenamiento	Conjunto de validación
Exactitud	0,87	0,60
Precisión	0,64	0,59
Sensibilidad	0,92	0,55

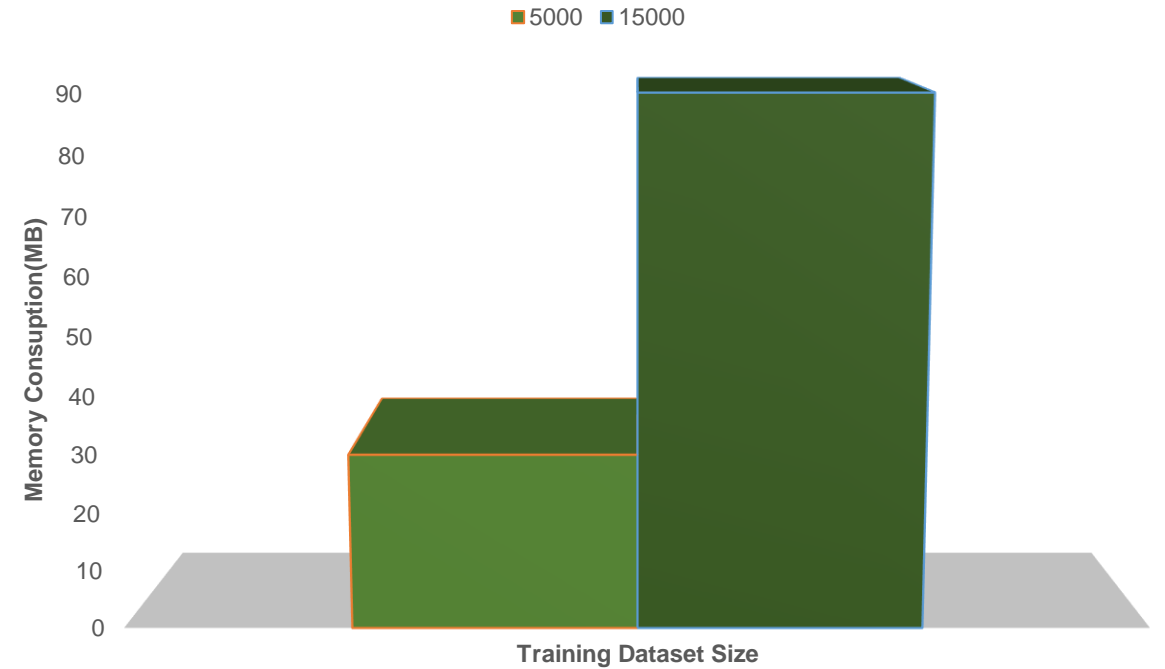
Métricas de evaluación obtenidas con el conjunto de datos de entrenamiento de 5,000 estudiantes y el conjunto de datos de validación de 15,000 estudiantes.



Consumo de tiempo y memoria




Consumo de tiempo



Consumo de memoria



C. Patiño-Forero, M. Agudelo-Toro, and M. Toro. Planning system for deliveries in Medellín. ArXiv e-prints, Nov. 2016. Available at: <https://arxiv.org/abs/1611.04156>

 Cornell University

arXiv.org > cs > arXiv:1611.04156

Computer Science > Data Structures and Algorithms

[Submitted on 13 Nov 2016]

Planning system for deliveries in Medellín

Catalina Patiño-Forero, Mateo Agudelo-Toro, Mauricio Toro

Here we present the implementation of an application capable of planning the shortest delivery route in the city of Medellín, Colombia. We discuss the different approaches to this problem which is similar to the famous Traveling Salesman Problem (TSP), but differs in the fact that, in our problem, we can visit each place (or vertex) more than once. Solving this problem is important since it would help people, especially stores with delivering services, to save time and money spent in fuel, because they can plan any route in an efficient way.

Comments: 5 pages, 9 figures

Subjects: **Data Structures and Algorithms (cs.DS)**

ACM classes: F.2.0; G.2.2

Cite as: [arXiv:1611.04156](https://arxiv.org/abs/1611.04156) [cs.DS]
(or [arXiv:1611.04156v1](https://arxiv.org/abs/1611.04156v1) [cs.DS] for this version)



¡GRACIAS!

