

Predicción
del éxito
académico,
en educación
superior,
utilizando
árboles de
decisión



Presentación del Equipo



Cristian
Alzate



Alejandro
Villada



Miguel
Correa



Mauricio
Toro



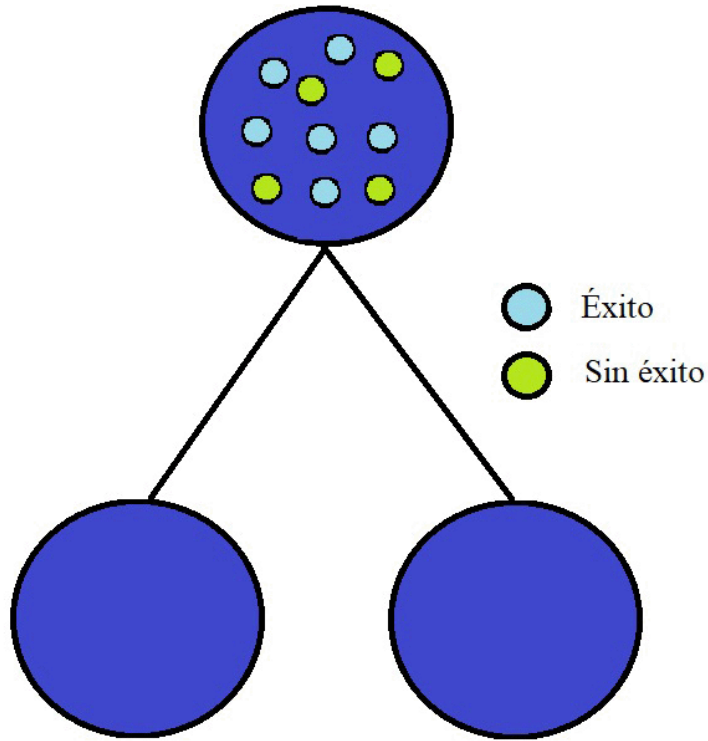
<http://github.com/avilladat/ST0245-001/tree/master/proyecto>



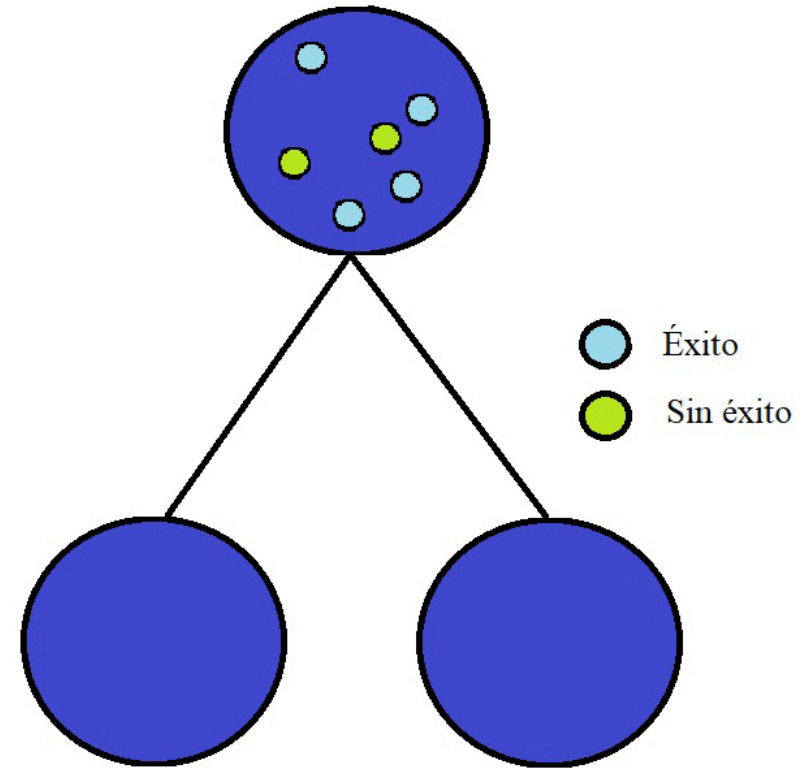


Entrenamiento del árbol de decisión binario ID3 para predecir los resultados de la prueba saber pro a partir de los resultados de la prueba saber 11°. En este ejemplo, mostramos un modelo con los puntajes de lenguaje y matemáticas.

División de un nodo



Esta división está basada en la condición “puntaje en lenguaje ≥ 4.5 .”
Para este caso, la impureza Gini de la izquierda es 0.44, la impureza Gini de la derecha es 0.5 y la impureza ponderada es de 0.458.



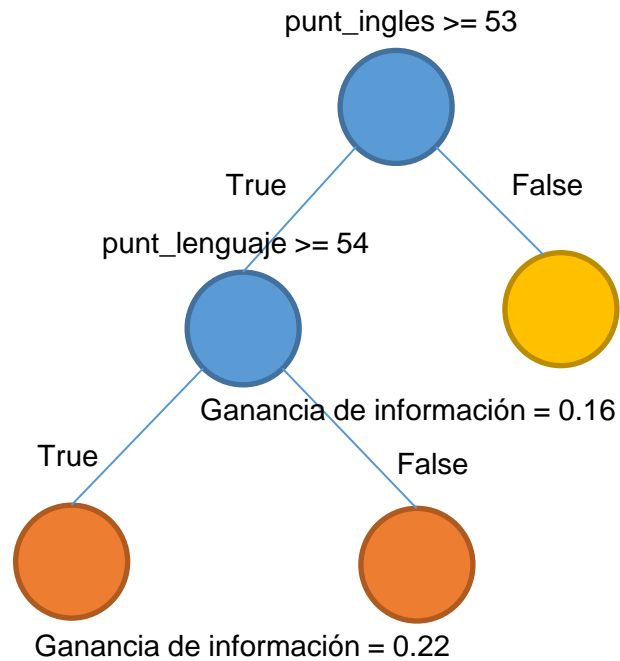
Esta división está basada en la condición “puntaje en matemática ≥ 4.5 .”
Para este caso, la impureza Gini de la izquierda es 0.375, la impureza Gini de la derecha es 0.5 y la impureza ponderada es 0.416.

	Complejidad en tiempo	Complejidad en memoria
Entrenamiento del modelo	$O(m^2 n \log n)$	$O(m^2 n)$
Validación del modelo	$O(n)$	$O(nm)$

Esta es la complejidad en memoria del algoritmo ID3 realizado. En esta caso m es el numero de atributos y n es el numero de datos.



Modelo de Árbol de Decisión



Un árbol de decisión ID3 para predecir el resultado del Saber Pro usando los resultados del Saber 11. Las mejores condiciones son las que tengan mayor ganancia de información. Los nodos amarillos son hojas con una baja ganancia de información, los naranjas con ganancia de información media y los rojos una ganancia de información alta

Características Más Relevantes



Puntaje biología



Puntaje Lenguaje



Puntaje Física



Puntaje Inglés



Puntaje Química



Numero de libros de la familia



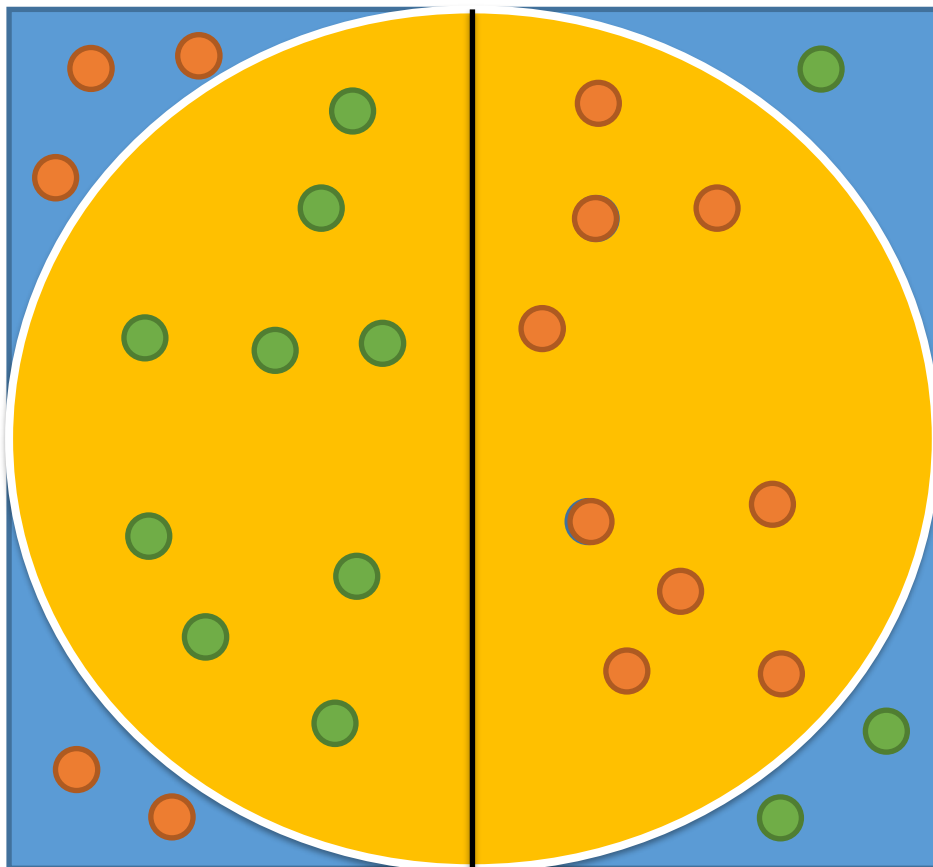
Puntaje Sociales



Puntaje matemáticas



Puntaje Filosofía

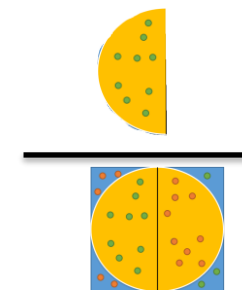


Modelo del árbol

La precisión es el número de predicciones correctas sobre el total de datos

La exactitud es el número de predicción correctas sobre el total de datos

La



Métricas de Evaluación

Conserven ese título

Completen esta lámina
en la tercera entrega



Crean la tabla en Powerpoint. ¡No copien
pantallazos pixelados del reporte, por
favor!

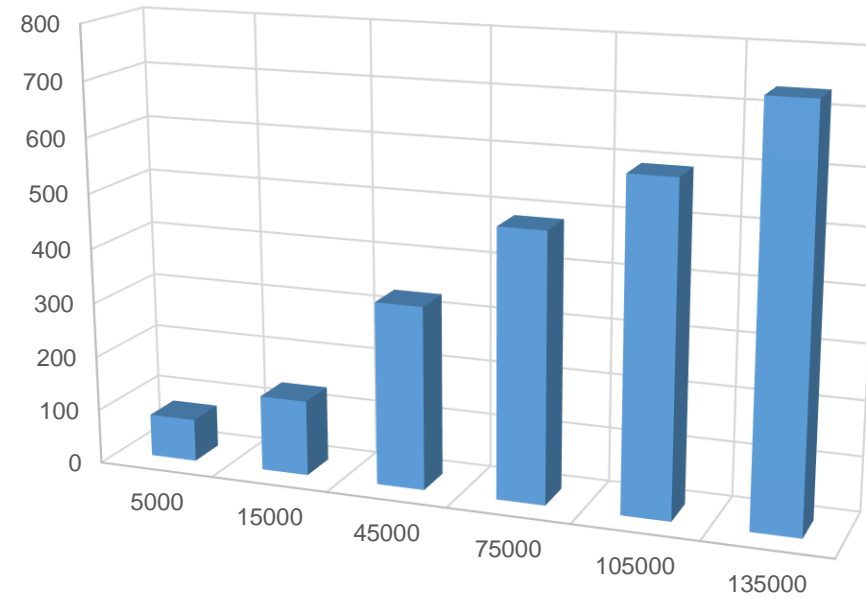
	Conjunto de entrenamiento	Conjunto de validación
Exactitud	0.8	0.62
Precisión	0.6	0.55
Sensibilidad	0.76	0.61

Métricas de evaluación obtenidas con el conjunto de datos de
entrenamiento de 135,000 estudiantes y el conjunto de datos
de validación de 45,000 estudiantes.



Incluyan otra gráfica en alta
definición relacionada con el
problema que están resolviendo.

Expliquen las tablas con sus
propias palabras



Consumo de tiempo

A close-up photograph of Barack Obama, looking down with his eyes closed, wiping his right eye with a white cloth. He is wearing a dark suit jacket and a white shirt. The background is dark and out of focus.

¡GRACIAS POR ESCUCHAR!

- https://www.google.com/search?q=icfes&tbm=isch&ved=2ahUKEwit3Zm7063sAhX8bDABHRQDAUwQ2-cCegQIABAA&oq=icfes&gs_lcp=CgNpbWcQAzIHCAAQsQMQQzICCAAYAggAMgIIADICCAAYAggAMgIIADICCAAYAggAMgIIAFCMUliMUmDpU2gAcAB4AIABtgGIAbYBkgEDMC4xmAEAoAEBqgELZ3dzLXdpei1pbWfAAQE&sclient=img&ei=so-DX628lvzZwbkPIlaE4AQ&bih=657&biw=1366#imgrc=TSLvRJ9FOXUFVM
- https://www.google.com/search?q=computadores+en+llamas&tbm=isch&ved=2ahUKEwj7pZ-uqlrtAhWrt1kKHejOCy4Q2-cCegQIABAA&oq=computadores+en+llamas&gs_lcp=CgNpbWcQAzoHCAAQsQMQQzoECAAAQzoCCAA6BQgAELEDOgYIABAIEB5QqKICWJuzAmDltQJoAHAAeACAAa8BiAGZDpIBBDAuMTOYAQCgAQGqAQtn3Mtd2l6LWltZ8ABAQ&sclient=img&ei=wCS0X_vFFKvv5gLona_wAg&bih=657&biw=1366#imgrc=RvtN9glt4jXhVM
- [https://www.google.com/search?q=lenguaje+icono&tbm=isch&ved=2ahUKEwj55sTurYrtAhUyrlkKHZK9AikQ2-cCegQIABAA&oq=lenguaje+icono&gs_lcp=CgNpbWcQAzICCAAYAggAMgIIADICCAAYAggAMgIIADICCAAYAggAMgIIADICCAAYAggAMgIIADICCAAYAggAMgIIADICCA6BggAEAcQHjoECAAQzoHCAAQsQMQQ1DsNVjuTGCOTmgAcAB4AYABugGIAAdMQkgEEMC4xNZgBAKABAaoBC2d3cy13aXotaW1nwAEB&sclient=img&ei=hSq0X7mQFbLc5gKS-4rIAG&bih=657&biw=1366#imgrc=PxoUKDwWzpvYvM](https://www.google.com/search?q=lenguaje+icono&tbm=isch&ved=2ahUKEwj55sTurYrtAhUyrlkKHZK9AikQ2-cCegQIABAA&oq=lenguaje+icono&gs_lcp=CgNpbWcQAzICCAAYAggAMgIIADICCAAYAggAMgIIADICCAAYAggAMgIIADICCAAYAggAMgIIADICCAAYAggAMgIIADICCAAYAggAMgIIADICCA6BggAEAcQHjoECAAQzoHCAAQsQMQQ1DsNVjuTGCOTmgAcAB4AYABugGIAAdMQkgEEMC4xNZgBAKABAaoBC2d3cy13aXotaW1nwAEB&sclient=img&ei=hSq0X7mQFbLc5gKS-4rIAG&bih=657&biw=1366#imgrc=PxoUKDwWzpvYvM)



- https://www.google.com/search?q=ingles+icono&tbm=isch&ved=2ahUKEwjwq6f0rIrtAhXPajABHcPtB6UQ2-cCegQIABAA&oq=ingles+icono&gs_lcp=CgNpbWcQAzICCAAyAggAMgIIADIGCAAQBxAeMgYIABAHEB4yCAgAEAcQBRAeMggIABAHEAUQHjIGCAAQBRAeMgYIABAFEB4yBggAEAUQHjoECAAQQ1DY6whYmfwIYNKACWgBcAB4AIAB0AGIAZEIkgEFMC42LjGYAQCgAQGqAQtn3Mtd2l6LWltZ8ABAQ&sclient=img&ei=hSm0X7BVz9XBuQ_D25-oCg&bih=657&biw=1366#imgrc=XHOeQeWg-PbeZM
- https://www.google.com/search?q=geografia+icono&tbm=isch&ved=2ahUKEwil6fzsrlrtAhV4bDABHUagD4sQ2-cCegQIABAA&oq=geografia+icono&gs_lcp=CgNpbWcQAzICCAAyAggAMgYIABAFEB4yBggAEAgQHjIGCAAQCBAeMgYIABAIEB46BggAEAcQHjoICAAQBxAFEb46CAgAEAgQBxAeUOxOWOlmyYP5raABwAHgAgAHZAYgBhgqSAQUwLjguMZgBAKABAaoBC2d3cy13aXotaW1nwAEB&sclient=img&ei=dSm0X-XaJvjYwbkPxsC-2Ag&bih=657&biw=1366#imgrc=BqqtH9qzq0rbTM