

Aplicación de la IA en educación:

PREDICCIÓN DEL RENDIMIENTO ACADÉMICO EN SECUNDARIA

Presentado por:

Mar Pérez

ÍNDICE

01 Introducción

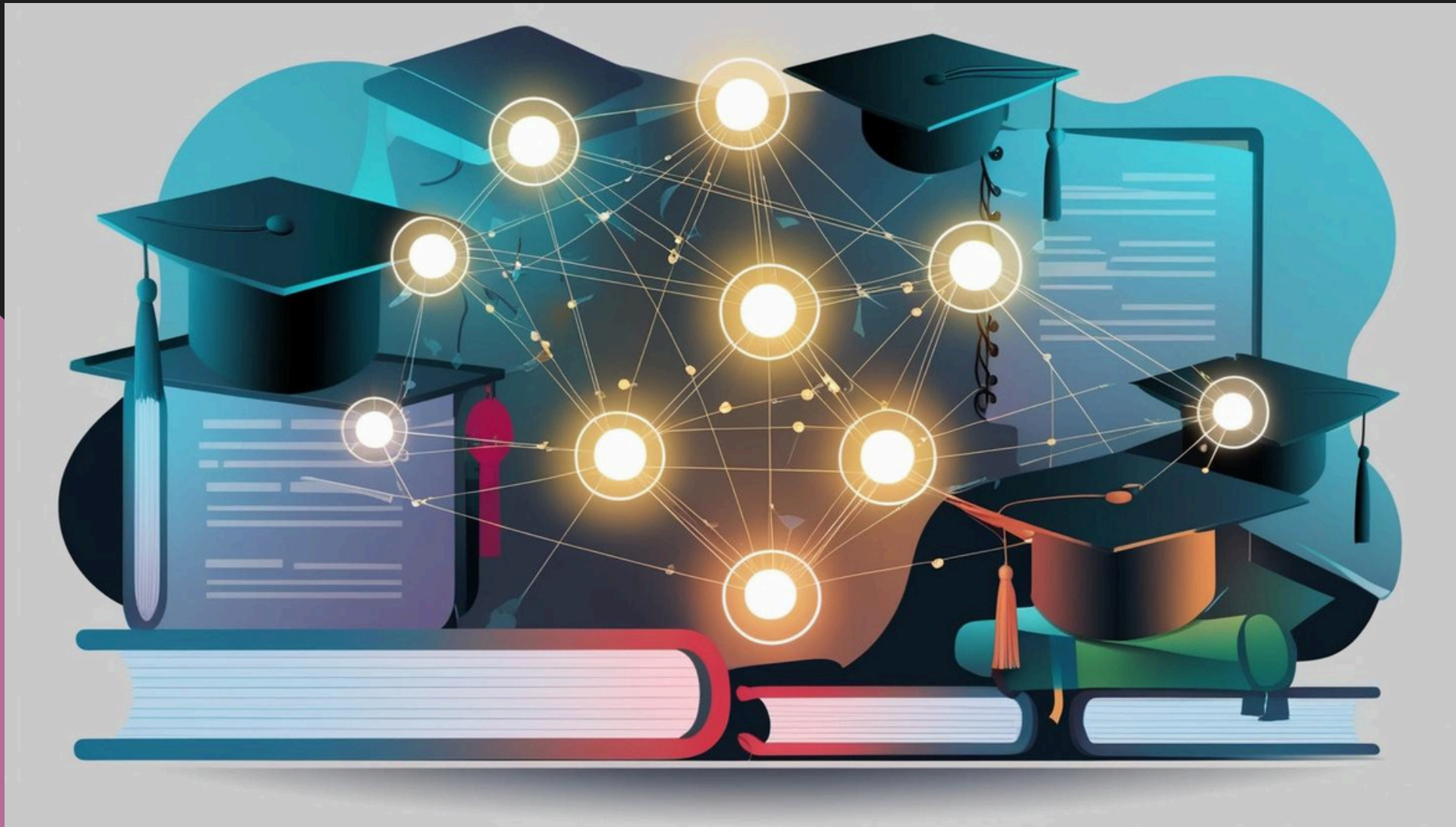
02 Datos y
preprocesamiento

03 Primer enfoque - Clasificación
multicategórica

04 Cambio de estrategia -
Conversión a binario

05 Resultados del nuevo
modelo

06 Conclusión



01. Introducción

- Creador: Rabie El Kharoua
- Tipo de problema: clasificación multicategórico.
- Objetivo: predecir las notas de los estudiantes de Secundaria.

0.2 Datos y preprocesamiento

- 2392 estudiantes
- 15 columnas
- No nulos
- No duplicados
- StudentID → índice
- Eliminar GPA → Data leakage





Variables

NUMÉRICAS

- StudyTimeWeekly
- Absences

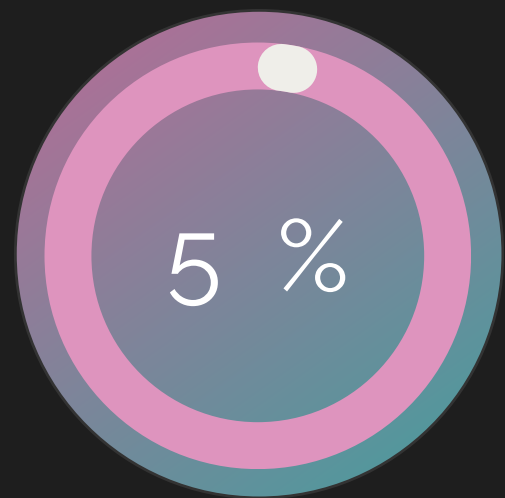
BINARIAS

- Gender
- Tutoring
- Extracurricular
- Sports
- Music
- Volunteering

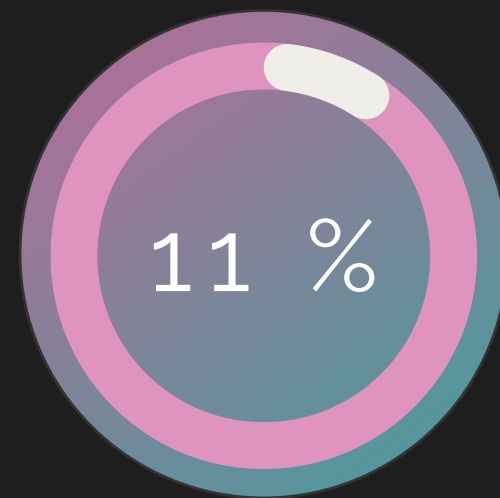
CATEGÓRICAS

- Age
- Ethnicity
- ParentalEducation
- ParentalSupport

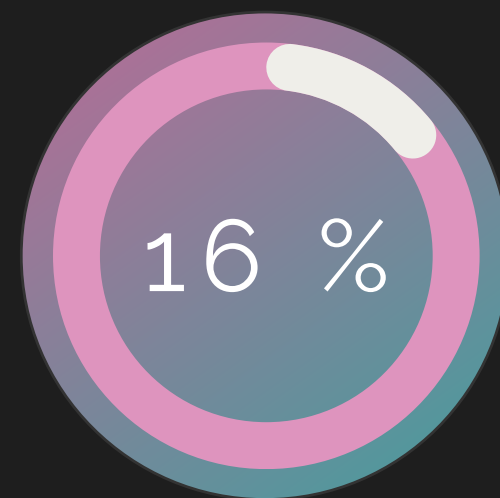
Distribución del target



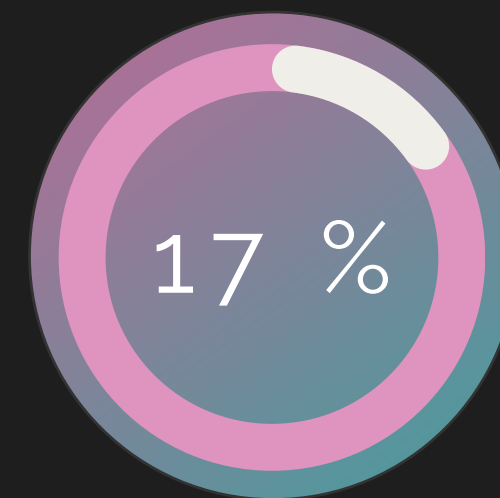
Clase 0



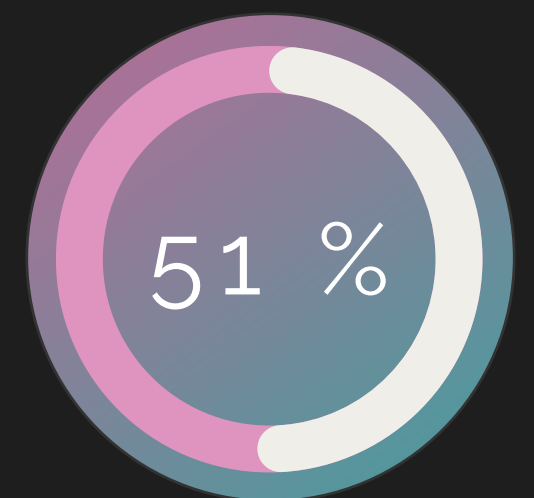
Clase 1



Clase 2



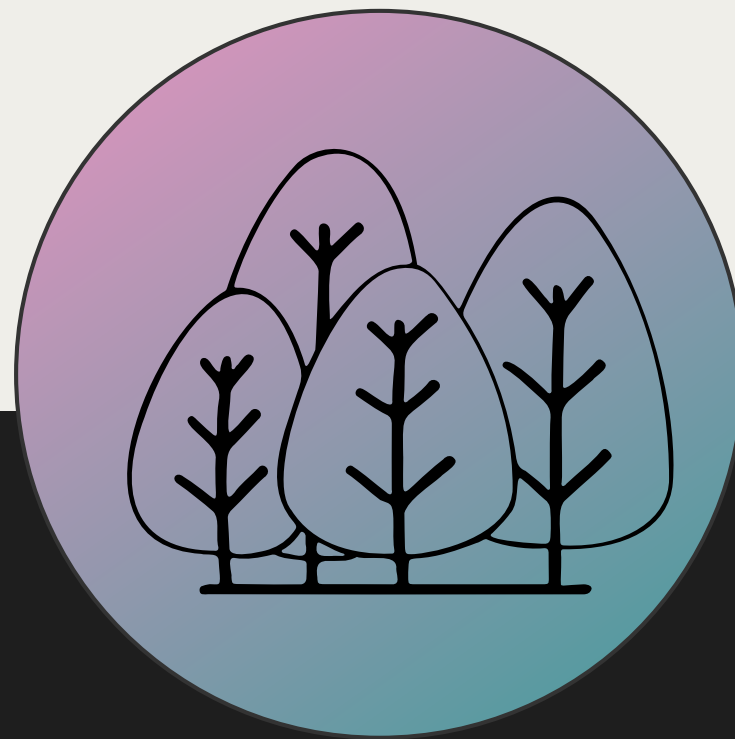
Clase 3



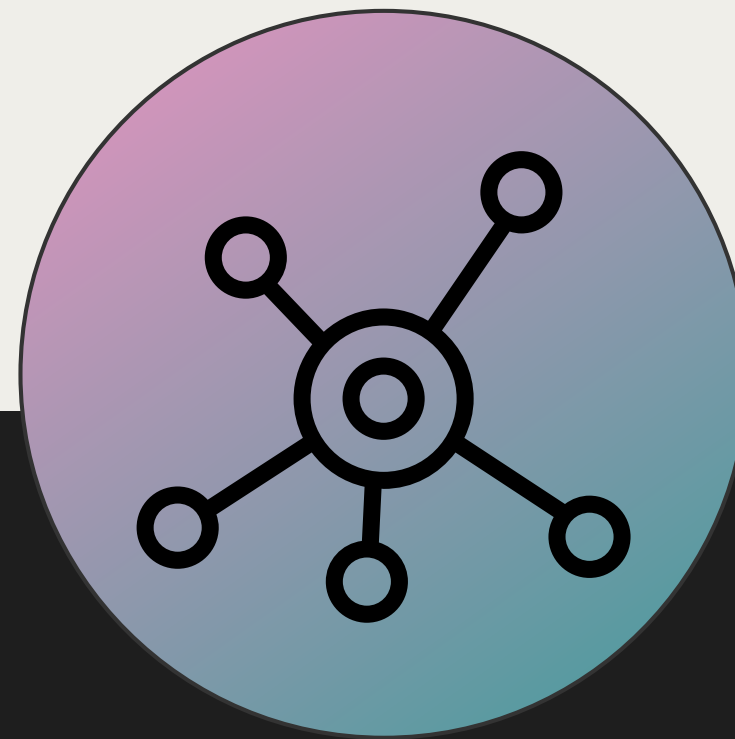
Clase 4

03. Primer enfoque - Clasificación multcategórica

Pipeline con modelos base:



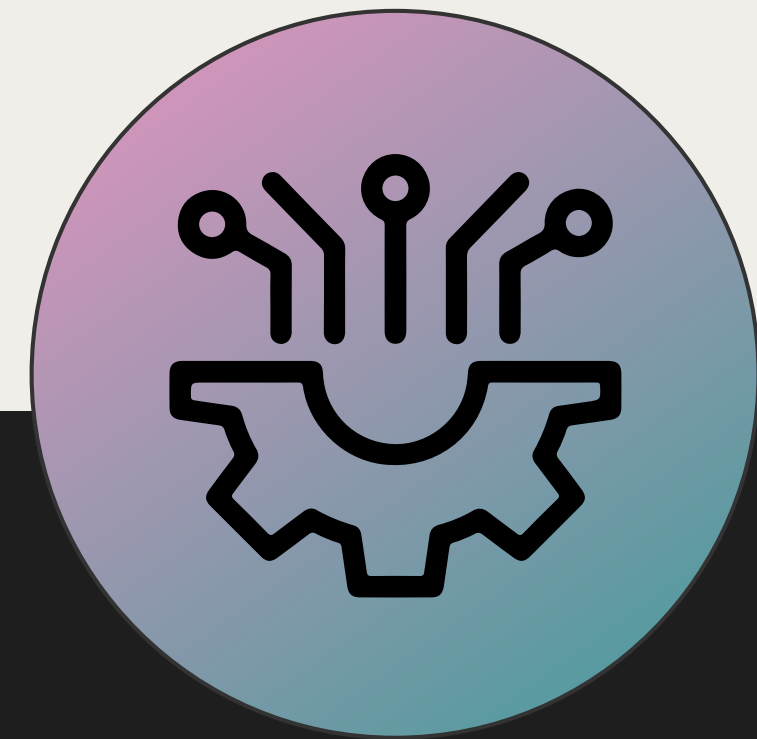
Random Forest



Gradient Boosting



LightGBM



XGBoost

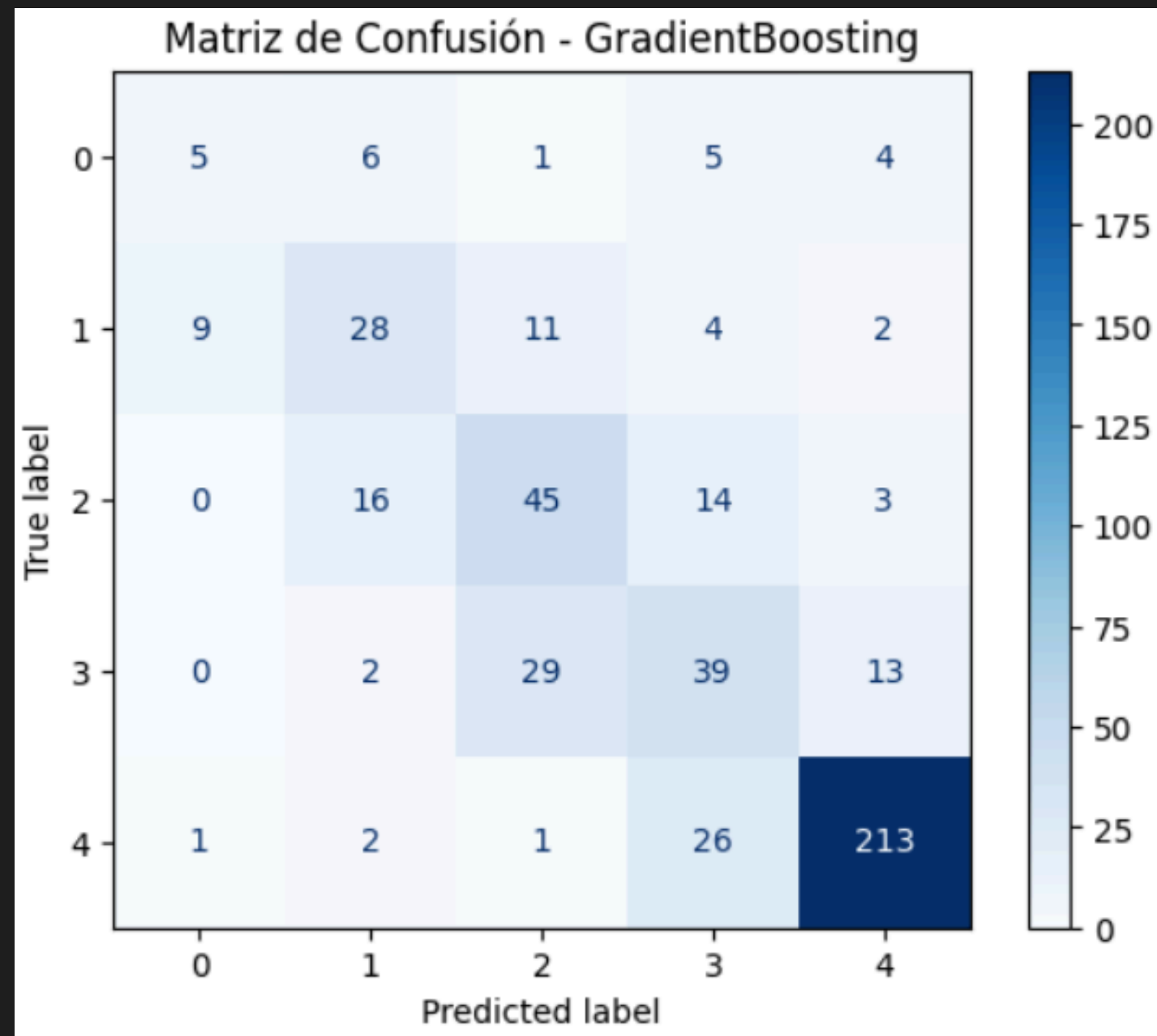


Mejor Modelo Optimizado

Gradient Boosting: Balance Accuracy - 0.5360

	Precision	Recall	F1 - Score	Support
0	0.33	0.24	0.28	21
1	0.52	0.52	0.52	54
2	0.52	0.58	0.55	78
3	0.44	0.47	0.46	83
4	0.91	0.88	0.89	243
Accuracy			0.69	479
Macro Avg	0.54	0.54	0.54	479
Weighted Avg	0.69	0.69	0.69	479

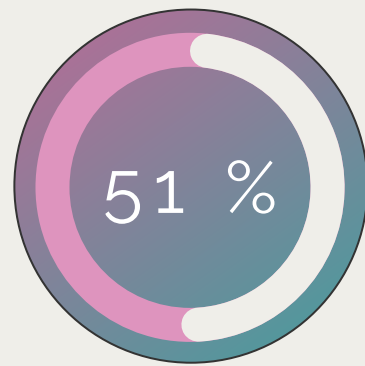
Matriz de Confusión



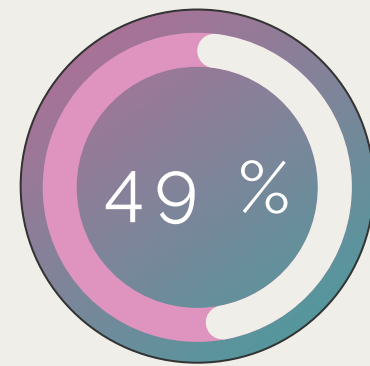
- Clase 0 y 1 → se predicen frecuentemente como otras clases
- Clase 3 → clase 2
- Clase 4 → Mejor predicha

04. Cambio de estrategia - Conversión a binario

- Nuevo Target:
 - Aprobado Vs Suspenso
 - Distribución:



Clase 0



Clase 1



Modelos



RANDOM FOREST



GRADIENT BOOSTING



LIGHTGBM



XGBOOST

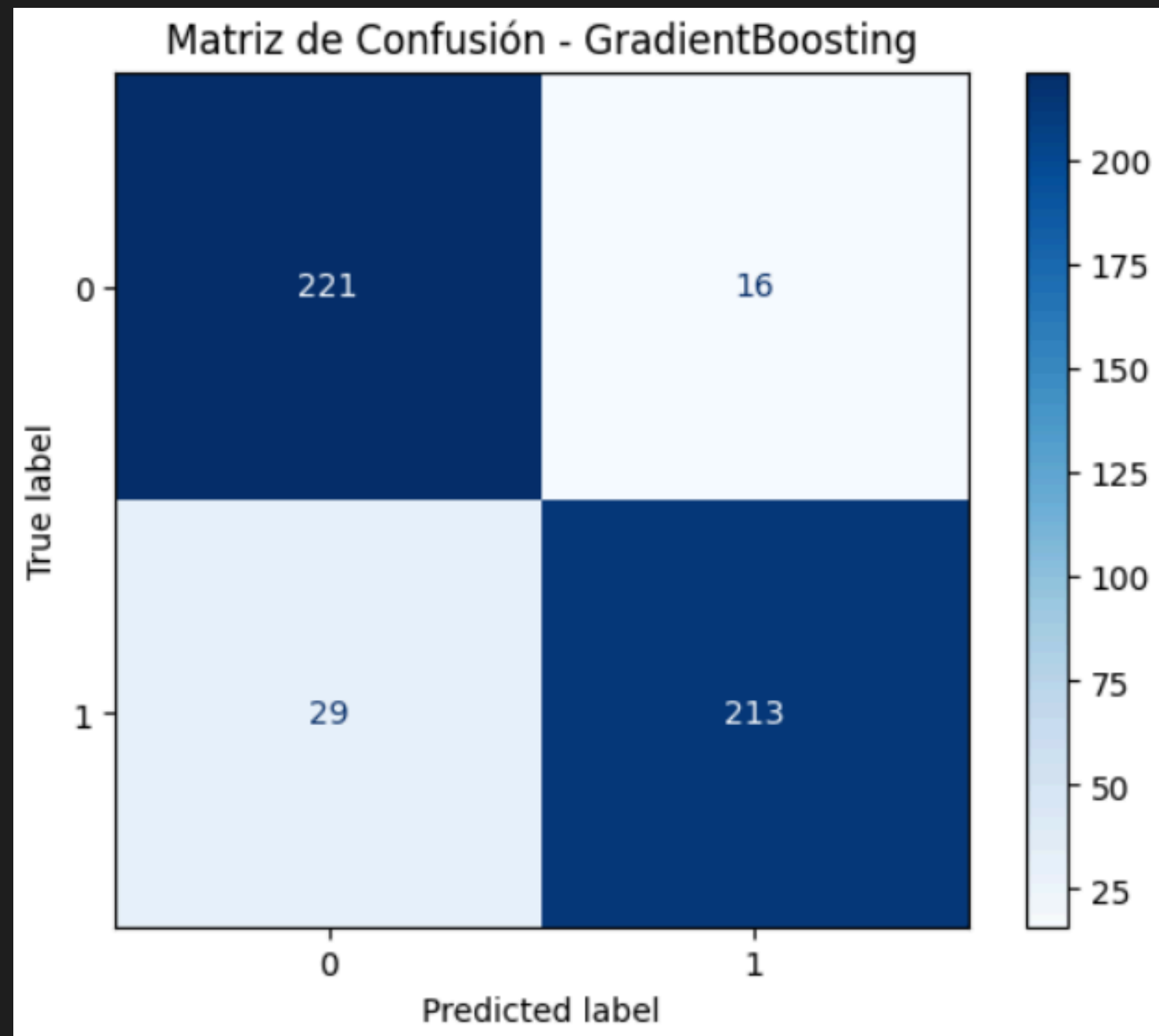


05. Resultados del nuevo modelo

Gradient Boosting: Balance Accuracy - 0.9063

	Precision	Recall	F1 - Score	Support
0	0.88	0.93	0.91	237
1	0.93	0.88	0.90	242
Accuracy			0.91	479
Macro Avg	0.91	0.91	0.91	479
Weighted Avg	0.91	0.91	0.91	479

Matriz de Confusión



- 01 CLASE 0
- 16 errores
 - 221 aciertos
- 02 CLASE 1
- 29 errores
 - 213 aciertos

06. Conclusiones

MEJORA RENDIMIENTO

- Robusto
- Eficiente

DISMINUCIÓN DE ERRORES

- Más concentrados
- Enfoque más claro

MAYOR PRECISIÓN Y RECALL

**MUCHAS
GRACIAS**