



Tecnológico de Monterrey

Actividad 4.3 (Regresión Logística Balanceo)

Lucero Jannete López García A01736938

Pilar Méndez Briones A01736843

Braulio González Esquivel

Ruben Flores Bello

Analítica de datos y herramientas de inteligencia artificial I (Gpo 101)

Fecha

02/04/2025

Caso 1	Botón-mini juego, juego, número de interacción			
Ponderación de clases '0' y '1'				
Exactitud	Precisión '0'	Precisión '1'	Sensibilidad '0'	Sensibilidad '1'
0.6346	0.6424	0.6226	0.42	0.81
Oversampling '0' '1'				
Exactitud	Precisión '0'	Precisión '1'	Sensibilidad '0'	Sensibilidad '1'
0.6579	0.6779	0.6349	0.4603	0.7980

Análisis: El modelo original tiene una exactitud moderada (0.63), lo que significa que aproximadamente el 63% de las predicciones totales fueron correctas. La precisión es mucho mayor para la clase 1 (0.81) que para la clase 0 (0.42), lo que indica que, cuando el modelo predice un 1, usualmente acierta, pero cuando predice un 0, comete muchos errores. La sensibilidad muestra que el modelo detecta mejor la clase positiva (0.81) que la negativa (0.42). Con oversampling mejora ligeramente la exactitud (0.66), y se equilibra un poco la detección entre ambas clases. Esto indica que el balance de clases ayudó a mejorar la identificación de la clase minoritaria.

Caso 2	Juego-'Usuario', 'mini juego', 'color presionado'			
Ponderación de clases '0' astro y '1' cadetes				
Exactitud	Precisión '0'	Precisión '1'	Sensibilidad '0'	Sensibilidad '1'
0.94	0.9934	0.8689	0.92	0.9854
Oversampling '0' '1'				
Exactitud	Precisión '0'	Precisión '1'	Sensibilidad '0'	Sensibilidad '1'
0.9494	0.9942	0.8831	0.9263	0.992

Análisis: Este es el modelo con mejor rendimiento general. Tiene una altísima exactitud (0.94), lo que significa que el modelo clasifica correctamente casi todos los casos. La precisión es sobresaliente tanto para la clase 0 (0.99) como para la clase 1 (0.86), lo que indica que ambas clases son bien distinguidas. Las sensibilidades también son altas, por lo que el modelo detecta con alta probabilidad cuando las clases están presentes. Con oversampling, mejora aún más (exactitud 0.949), reafirmando que es un modelo muy confiable para predecir el tipo de juego.

Caso 3	Auto push- 'Usuario', 'mini juego', 'color presionado'			
---------------	---	--	--	--

Ponderación de clases '0' astro y '1' cadetes				
Exactitud	Precisión '0'	Precisión '1'	Sensibilidad '0'	Sensibilidad '1'
0.7664	0.9977	0.03854	0.7660	0.7826
Oversampling '0' '1'				
Exactitud	Precisión '0'	Precisión '1'	Sensibilidad '0'	Sensibilidad '1'
0.7647	0.9970	0.06	0.7645	0.81

Análisis: Aunque la exactitud es buena (0.76), hay un gran desequilibrio en la precisión: 0.997 para clase 0 y apenas 0.038 para clase 1. Esto significa que el modelo rara vez acierta cuando predice la clase 1. La sensibilidad para ambas clases es similar (alrededor de 0.78), pero el bajo valor de precisión en la clase 1 indica que hay muchos falsos positivos. Con oversampling, la sensibilidad de la clase 1 mejora levemente (de 0.78 a 0.81), pero la precisión sigue siendo muy baja. El modelo requiere ajustes adicionales, como incorporar nuevas variables o aplicar regularización.

Caso 4	Usuario- 'tiempo de interacción', 'botón correcto', 'número de interacción'			
Ponderación de clases ‘valentin’ ‘demás usuarios’				
Exactitud	Precisión ‘demás usuarios’	Precisión ‘Valentin’	Sensibilidad ‘demás usuarios’	Sensibilidad ‘valentin’”
0.4625	0.9823	0.03	0.4559	0.90
Oversampling ‘valentin’ ‘demás usuarios’				
Exactitud	Precisión ‘demás usuarios’	Precisión ‘Valentin’	Sensibilidad ‘demás usuarios’	Sensibilidad ‘valentin’”
0.4738	0.9805	0.0382	0.4696	0.6923

Análisis: El modelo es ineficiente en predecir a Valentín. Aunque la sensibilidad es alta (0.90), lo que indica que el modelo detecta la mayoría de los casos reales de Valentín, la precisión es muy baja (0.03), lo que significa que casi todos los casos predichos como Valentín son en realidad falsos positivos. Con oversampling, la precisión mejora ligeramente (0.038) pero la sensibilidad disminuye (0.69). Esto refleja un problema de sobreajuste o ruido en los datos.

Caso 5	Usuario- 'tiempo de interacción', 'botón correcto', 'número de interacción'
--------	---

Ponderación de clases 'Yael david' 'demás usuarios'				
Exactitud	Precisión 'demás usuarios'	Precisión 'Yael david'	Sensibilidad 'demás usuarios'	Sensibilidad 'Yael david'
0.6068	0.9801	0.0478	0.6066	0.6101
Oversampling 'Yael david' 'demás usuarios'				
Exactitud	Precisión 'demás usuarios'	Precisión 'Yael david'	Sensibilidad 'demás usuarios'	Sensibilidad 'Yael david'
0.6039	0.9837	0.0460	0.6018	0.6458

Análisis: Similar al caso anterior, el modelo tiene baja precisión para Yael David (0.047), aunque la sensibilidad es más equilibrada (0.61). Esto indica que el modelo detecta algunos casos de Yael, pero también se equivoca mucho. Con oversampling, la sensibilidad mejora a 0.645 pero la precisión baja aún más (0.046). Este comportamiento sugiere que el modelo puede estar generando muchas falsas alarmas al intentar detectar a este usuario.

Caso 6	Dificultad- 'tiempo de interacción', 'mini juego', 'juego'			
Ponderación de clases 'Nivel básico' 'Nivel avanzado'				
Exactitud	Precisión 'Nivel básico'	Precisión 'Nivel avanzado'	Sensibilidad 'Nivel básico'	Sensibilidad 'Nivel avanzado'
0.533	0.688	0.3688	0.5335	0.5276
Oversampling 'Nivel básico' 'Nivel avanzado'				
Exactitud	Precisión 'Nivel básico'	Precisión 'Nivel avanzado'	Sensibilidad 'Nivel básico'	Sensibilidad 'Nivel avanzado'
0.535	0.691	0.409	0.535	0.576

Análisis: El modelo tiene un rendimiento bajo. La exactitud está en 0.53, lo que está apenas por encima del azar, y las precisiones y sensibilidades para ambas clases son moderadas pero bajas, lo que indica que el modelo no está logrando distinguir adecuadamente entre los niveles de dificultad. El oversampling mejora levemente la sensibilidad para la clase avanzada, pero no es suficiente para hacer al modelo confiable. Se recomienda revisar la calidad y relevancia de las variables predictoras.

Caso 7	Colores- 'tiempo de interacción', 'botón correcto', 'número de interacción'
Ponderación de clases 'Violeta' 'demás colores'	

Exactitud	Precisión 'demás colores	Precisión 'Violeta'	Sensibilidad 'demás colores	Sensibilidad 'Violeta'
0.646	0.752	0.354	0.745	0.327
Oversampling 'Violeta' 'demás colores				
Exactitud	Precisión 'demás colores	Precisión 'Violeta'	Sensibilidad 'demás colores	Sensibilidad 'Violeta'
0.635	0.752	0.360	0.743	0.295

Análisis: El modelo tiene una exactitud razonable (0.646), pero al igual que en casos anteriores con clases desbalanceadas, su precisión para la clase minoritaria (Violeta) es baja (0.354) y la sensibilidad también (0.327), lo que implica que predice pocos casos de Violeta y muchos de ellos son erróneos. Con oversampling, la precisión mejora mínimamente y la sensibilidad disminuye. Es probable que se necesiten nuevas características que mejor representen el color presionado.

Caso 8	Colores- 'tiempo de interacción', 'botón correcto', 'número de interacción'			
Ponderación de clases 'Green' 'demás colores'				
Exactitud	Precisión 'demás colores'	Precisión Verde	Sensibilidad 'demás colores'	Sensibilidad Verde
0.376	0.794	0.247	0.412	0.741
Oversampling 'Green' 'demás colores'				
Exactitud	Precisión 'demás colores'	Precisión Verde	Sensibilidad 'demás colores'	Sensibilidad Verde
0.430	0.782	0.248	0.326	0.739

Análisis: Este modelo tiene una exactitud baja (0.37), y aunque la sensibilidad para Verde es alta (0.74), la precisión es muy baja (0.24), indicando que aunque detecta muchos verdes, se equivoca mucho en sus predicciones. Con oversampling, la exactitud mejora a 0.43, pero la precisión permanece baja. El modelo está sobreajustado a la clase verde sin poder identificarla con fiabilidad.

Caso 9	Mini juego- 'tiempo de interacción', 'botón correcto', 'Juego'			
Ponderación de clases 'Despegue' 'Mini Juego tipo A'				
Exactitud	Precisión	Precisión 'Mini	Sensibilidad	Sensibilidad

	'Despegue'	Juego tipo A'	'Despegue'	'Mini Juego tipo A'
0.646	0.342	0.985	0.963	0.572
Oversampling 'Despegue' 'Mini Juego tipo A'				
Exactitud	Precisión 'Despegue'	Precisión 'Mini Juego tipo A'	Sensibilidad 'Despegue'	Sensibilidad 'Mini Juego tipo A'
0.611	0.306	0.985	0.942	0.573

Análisis: Este modelo es aceptable. Tiene una exactitud de 0.64 y alta sensibilidad para ambas clases, especialmente para Despegue (0.96), aunque su precisión es baja (0.34). Esto implica que detecta bien la clase Despegue, pero con muchas predicciones incorrectas. Con oversampling, hay una pequeña caída en la exactitud pero la sensibilidad se mantiene estable. Es un modelo prometedor con posibilidad de mejora.

Caso 10	Número de interacciones- 'tiempo de interacción', 'color presionado'			
Ponderación de clases 'Interacciones del 1 al 128', 'Interacciones del 129 al 256'				
Exactitud	Precisión 'Interacciones del 1 al 128'	Precisión 'Interacciones del 129 al 256'	Sensibilidad 'Interacciones del 1 al 128'	Sensibilidad 'Interacciones del 129 al 256'
0.928	0.998	0.211	0.928	0.942
Oversampling 'Interacciones del 1 al 128', 'Interacciones del 129 al 256'				
Exactitud	Precisión 'Interacciones del 1 al 128'	Precisión 'Interacciones del 129 al 256'	Sensibilidad 'Interacciones del 1 al 128'	Sensibilidad 'Interacciones del 129 al 256'
0.931	1.0	0.217	0.942	0.6458

Análisis: Este modelo tiene el mejor balance general. La exactitud es alta (0.93), la precisión para ambas clases es muy buena (0.998 y 0.211), y la sensibilidad también está bien equilibrada. Con oversampling, los resultados mejoran ligeramente, alcanzando precisión perfecta (1.0) para la clase mayoritaria y aceptable para la clase minoritaria. Es un modelo robusto y confiable para categorizar por niveles de interacción.