



# Tecnológico de Monterrey

**Instituto Tecnológico y de Estudios Superiores de Monterrey Campus Puebla**

## **Actividad 4.3 (Regresión Logística Ajuste)**

### **Equipo 5:**

José Manuel Morales Escalante

Josué Ibáñez Santana

Sergio Giles Bracamontes

Ismael Martinez Duran

### **Materia:**

Analítica de datos y herramientas de inteligencia artificial II

### **Fecha:**

27 de abril de 2025

## Análisis Profundo de Hallazgos: Comparación entre Reponderación y Oversampling

El análisis detallado de la tabla revela diferencias clave entre los modelos entrenados con **Reponderación** y **Oversampling**, aunque en la mayoría de los casos los resultados son muy similares. Sin embargo, algunos modelos presentan variaciones significativas que merecen atención.

### 1. Modelos con Diferencias Notables entre Técnicas de Balanceo

#### Autopush

- **Precisión Label 1:**
  - Reponderación: **0.990272**
  - Oversampling: **0.991228** (ligera mejora)
- **Sensibilidad Label 1:**
  - Reponderación: **0.335089**
  - Oversampling: **0.371955** (mejoró aproximadamente 3.7%)
- **Exactitud ("Execitual"):**
  - Reponderación: **0.337035**
  - Oversampling: **0.373612** (mejoró aproximadamente 3.7%)
- **Conclusión:**
  - Oversampling logra una mejora marginal pero consistente en todas las métricas.
  - La sensibilidad sigue siendo baja, lo que sugiere que el modelo sigue teniendo dificultades para detectar casos positivos.

#### Color Rojo

- **Precisión Label 1:**
  - Reponderación: **0.014063**
  - Oversampling: **0.014634** (ligero aumento)
- **Exactitud ("Execitual"):**

- Reponderación: **0.587851**
- Oversampling: **0.604180** (mejoró aproximadamente 1.6%)
- **Conclusión:**
  - Aunque la mejora es pequeña, Oversampling consigue un mejor equilibrio general.
  - La sensibilidad perfecta (1.0) con precisión extremadamente baja sugiere un sobreajuste hacia la clase mayoritaria.

## Color Violeta

- **Precisión Label 1:**
  - Reponderación: **0.366972**
  - Oversampling: **0.375415** (mejoró aproximadamente 2.3%)
- **Sensibilidad Label 1:**
  - Reponderación: **0.300000**
  - Oversampling: **0.282500** (empeoró aproximadamente 5.8%)
- **Exactitud ("Execitua"):**
  - Reponderación: **0.681907**
  - Oversampling: **0.689745** (mejoró aproximadamente 1.1%)
- **Conclusión:**
  - Oversampling aumenta la precisión pero reduce la sensibilidad, lo que indica un trade-off.
  - La exactitud mejora ligeramente, pero el modelo sigue siendo poco sensible para Label 1.

## Color Verde

- **Sensibilidad Label 1:**
  - Reponderación: **0.744186**
  - Oversampling: **0.759690** (mejoró aproximadamente 2.1%)
- **Exactitud ("Execitua"):**
  - Reponderación: **0.387329**

- Oversampling: **0.377531** (empeoró aproximadamente 2.5%)
- **Conclusión:**
  - Oversampling mejora la detección de Label 1, pero empeora el equilibrio general.
  - Esto podría deberse a un sobreajuste en la clase minoritaria.

## Minijuego Asteroides

- **Sensibilidad Label 1:**
  - Reponderación: **0.703846**
  - Oversampling: **0.734615** (mejoró aproximadamente 4.4%)
- **Exactitud ("Execitua"):**
  - Reponderación: **0.335728**
  - Oversampling: **0.306336** (empeoró aproximadamente 8.7%)
- **Conclusión:**
  - Oversampling aumenta la sensibilidad pero reduce la exactitud, lo que sugiere que se están introduciendo más falsos positivos.

## 2. Modelos con Resultados Prácticamente Idénticos

Algunos modelos no muestran diferencias significativas entre técnicas de balanceo, lo que sugiere que el oversampling no aporta mejoras en estos casos:

- **Botón correcto** (igual en todas las métricas).
- **Color azul** (ningún cambio).
- **Color amarillo** (diferencias mínimas, menos del 1%).
- **Minijuego restaurantes** (diferencias insignificantes).

### 3. Hallazgos Clave

#### Problemas Comunes:

- **Trade-off entre precisión y sensibilidad:**
    - En modelos como **color violeta** y **minijuego asteroides**, aumentar la sensibilidad reduce la exactitud, lo que indica que el modelo está clasificando más ejemplos como positivos pero con menor certeza.
  - **Métricas extremas en ciertos modelos:**
    - **Autopush** tiene precisión casi perfecta (0.99) pero sensibilidad muy baja (aproximadamente 0.35), lo que sugiere que solo predice positivos cuando está muy seguro, pero falla en detectar muchos casos reales.
    - **Color rojo** tiene sensibilidad perfecta (1.0) pero precisión bajísima (0.014), lo que implica que casi todas las predicciones positivas son incorrectas (alto número de falsos positivos).
  - **Impacto limitado del oversampling:**
    - En la mayoría de los casos, no hay una mejora clara al usar oversampling en lugar de reponderación.
    - Solo en **autopush, color rojo y violeta** hay diferencias ligeramente favorables al oversampling.
- 

#### Conclusiones Finales y Recomendaciones

1. **Priorizar Reponderación en modelos sin diferencias** (botón correcto, color azul, amarillo, minijuego restaurantes), ya que oversampling no aporta beneficios.
2. **Evaluar técnicas alternativas** (como SMOTE o ajustes de umbral) en modelos con métricas extremas (autopush, color rojo).
3. **Analizar el trade-off en modelos como color violeta y minijuego asteroides:** ¿Es preferible mayor sensibilidad aunque baje la exactitud?
4. **Revisar el desbalanceo de clases**, ya que algunos modelos (color rojo) tienen sensibilidad perfecta pero precisión casi nula, lo que sugiere un problema de distribución de clases.

**Acción recomendada:** Realizar un análisis de curvas ROC y matriz de confusión para entender mejor los falsos positivos/negativos en los modelos más críticos.