

Machine Learning E-Commerce



Raquel Hernández Lozano
The Bridge 2025

Contextualización Ecommerce

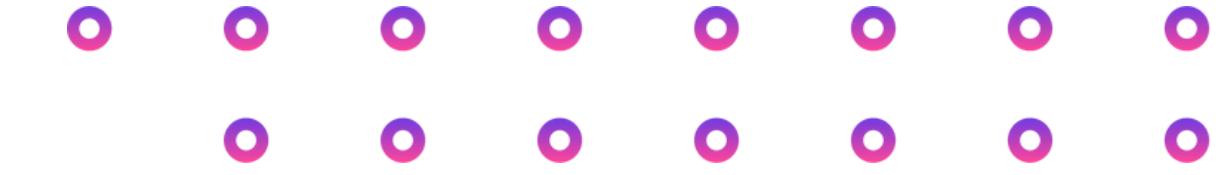
NEW!





Problema de Negocio





Problema de Negocio

La mayoría del tráfico web no convierte.

Necesitamos identificar a los clientes reales antes de que se vayan

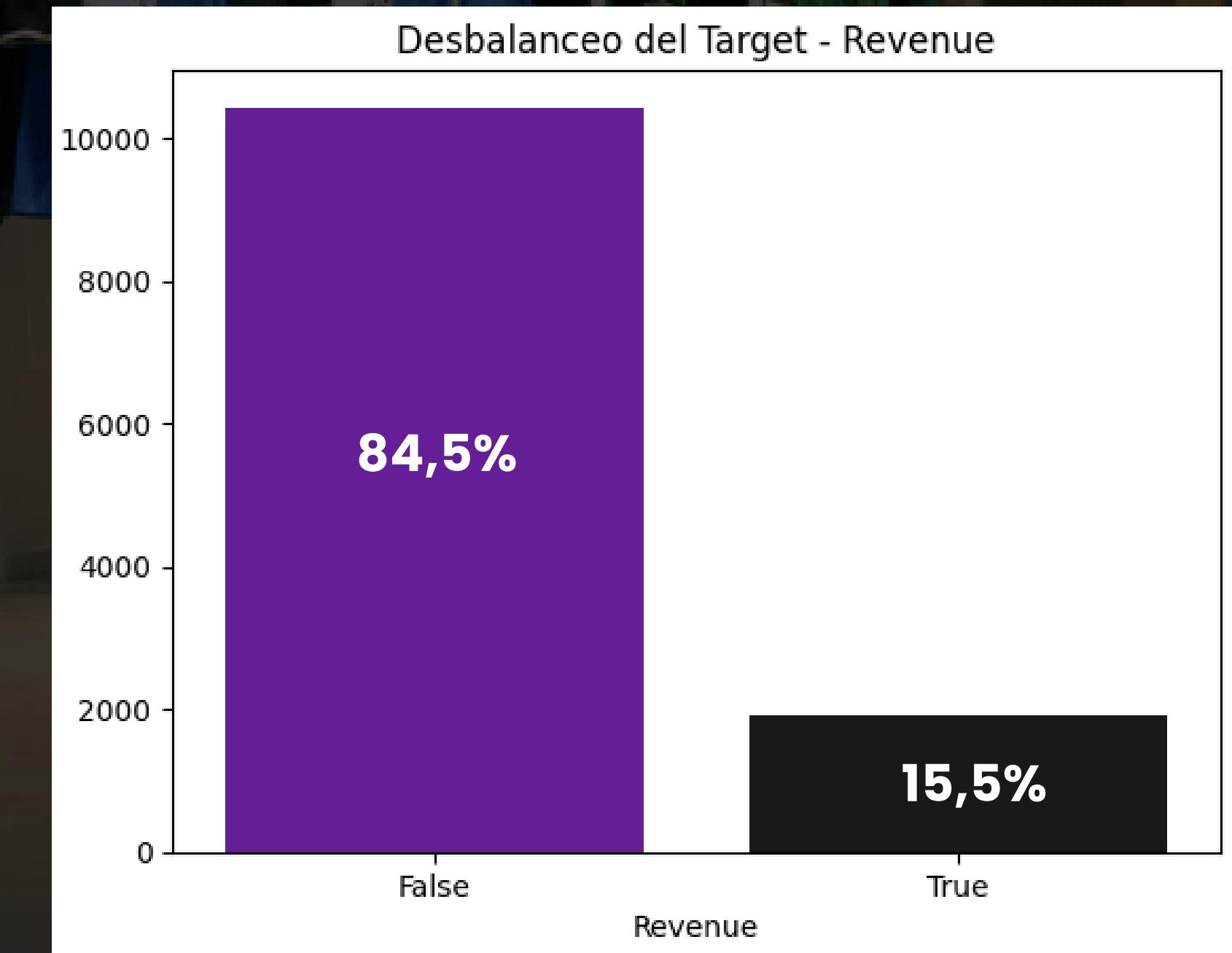


Reto E- Commerce:

La mayoría de sesiones no termina en compra.

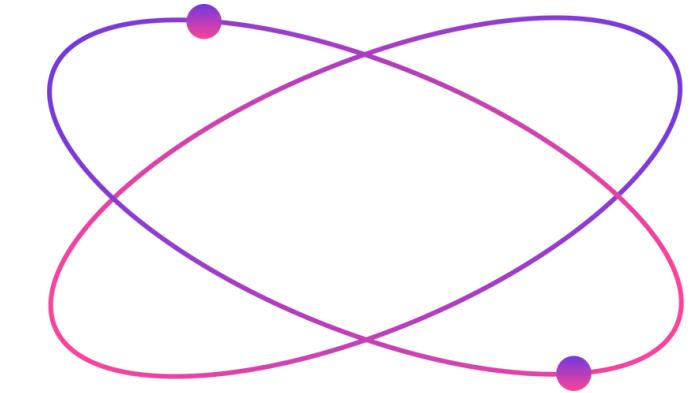
Desbalanceo de clases

Total de sesiones: 12.330
18 variables



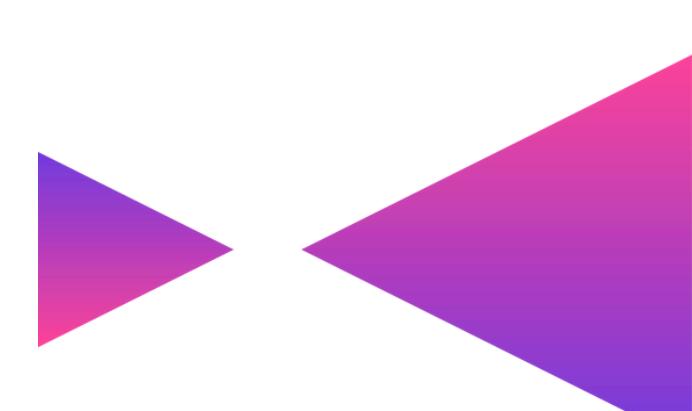


Fuente de datos disponible



Google Analytics

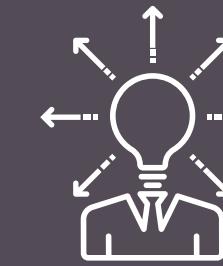
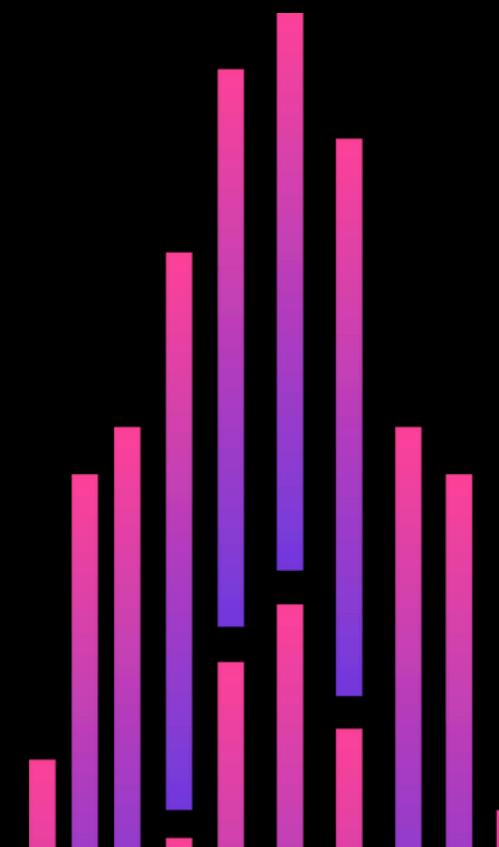
GA4 sera nuestro combustible para que el modelo de machine learning funcione correctamente para predecir si la sesion va a tener compra o no



Huella digital de clientes



Google Analytics



Variables de Comportamiento Web

Administrative_Duration
Informational_Duration
ProductRelated_Duration

ExitRates
PageValues ***



Variables Contextuales

Browser
Region
TrafficType
VisitorType

Month
SpecialDay
Weekend
....



Variable Objetivo

Revenue YES/NO

PageValue info GA

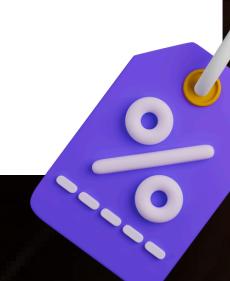
APLICANDO MODELO MACHINE LEARNING: CLASIFICACIÓN

Predecir la variable target (compra/no compra) a nivel de sesión

**Mejorar previsiones
de producción para
evitar futuras
rotura de stock**

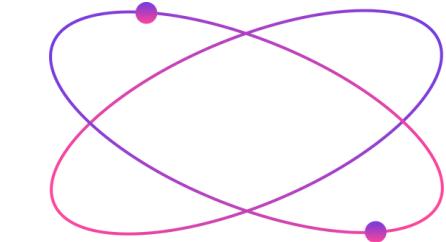


**Mejorar en ventas
con estrategia de
marketing bien
enfocada**





La "Magia" del Machine Learning



Modelado Predictivo

"Nos interesa que el modelo no se pierda ninguna posible compra. Si hay una compra, queremos que el modelo la detecte"

Métrica

Recall Alto



Buscamos crear un detector de clientes en nuestra web.

**Modelo seleccionado y optimizado SVM:
Support Vector Maching**

Algoritmos Machine Learning
Supervisados

Algoritmo Machine Learning **No**
Supervisado



REAL TIME: PROBEMOS EL MODELO

Streamlit

Modelo de Predicción de Intención de Compra (E-commerce)

Ajusta las variables clave para predecir la intención de compra. Ahora con manejo robusto de tipos de datos para evitar errores.

Parámetros de la Sesión

Variable	Tipo	Valor Actual
PageValues (Valor Promedio)	Slider	18.00
ProductRelated_Duration (Duración en Prod.)	Slider	600.00
Administrative_Duration (Duración en Admin.)	Slider	7.50
Informational_Duration (Duración en Info.)	Slider	48.90

Predecir Intención de Compra

Duración en Servicios

Contexto del Usuario

¿La sesión ocurrió durante el Fin de Semana?

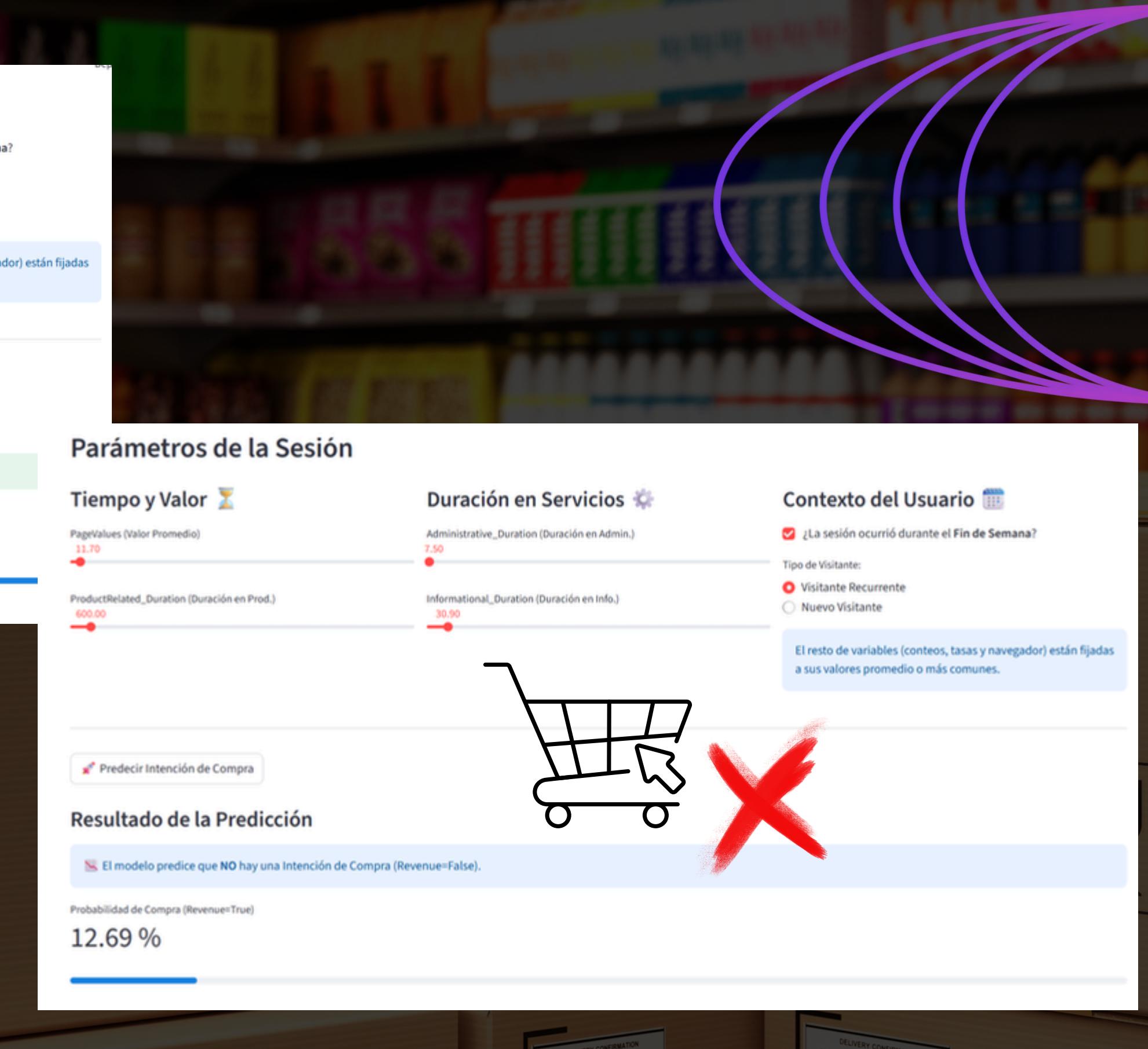
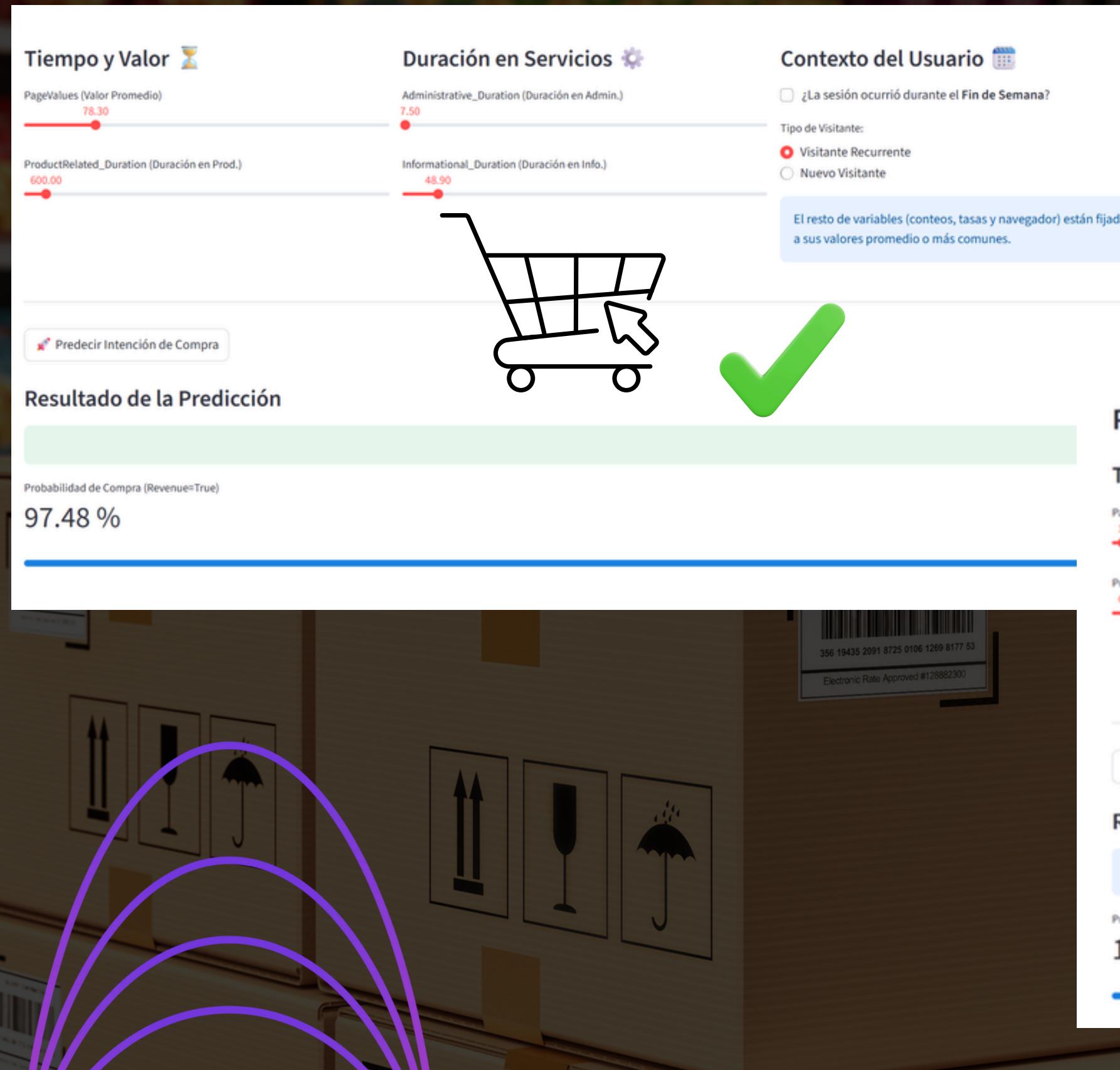
Tipo de Visitante:

Visitante Recurrente

Nuevo Visitante

El resto de variables (conteos, tasas y navegador) están fijadas a sus valores promedio o más comunes.

PRECICCIONES MODELO SVM OPTIMIZADO



Limitaciones

Datos solo de 1 año.

Falta automatización. Futuro csv para subida de datos.

Seguir alimentando el modelo, para mejorar futuras clasificaciones.

Escalabilidad

Posibilidad de reentrenar el modelo con más datos.

Si el tráfico se duplica, duplicamos aprendizaje.

No afecta en nada a la velocidad de carga de web.

Posibilidad de enriquecer el modelo incrementando la precisión.

Incluir datos offline. Clientes registrados.

Futuros beneficios de usar Machine Learning para predecir clientes

01

Aumento la tasa de conversión %.
Código dto/chatbot.



02

Maximización ROI en Marketing.
Publicidad dirigida y bien enfocada.
Gasto publicitario transformado en
inversión inteligente



03

Optimización de customer
Journey mejorando experiencia
de cliente.



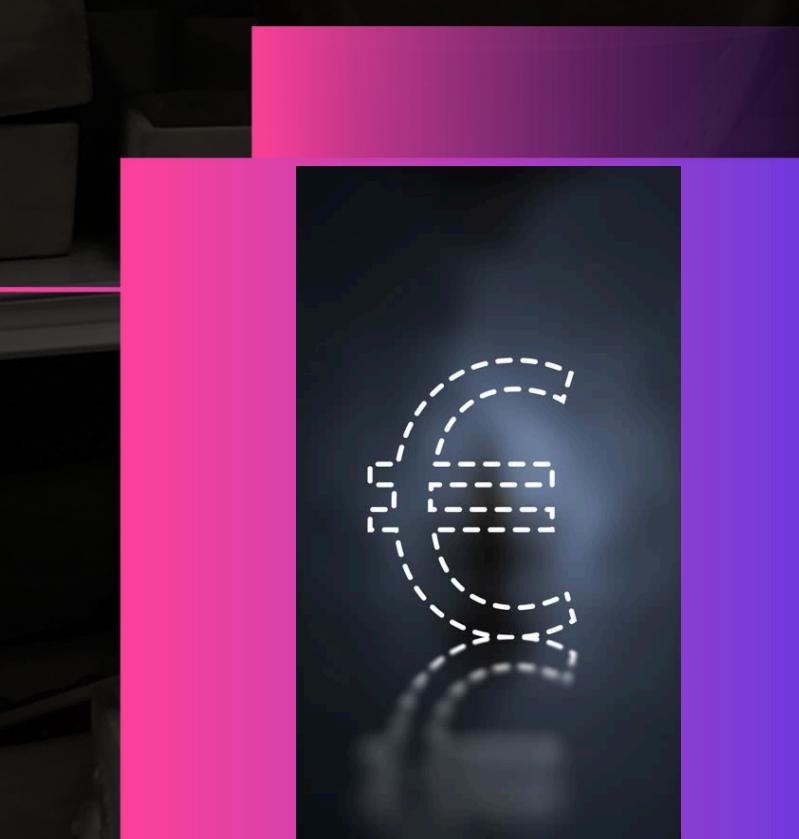
04

Mejorar prevision de stock con
las señales de intención de
compra.



05

El modelo de ML aumenta la
inteligencia del negocio.
Predicción del valor de vida de cliente





¡Gracias por vuestra atención!