



Aplicación Estratégica de Modelos Predictivos e Insights en eCommerce

1. Aplicación de Insights para Decisiones Estratégicas

A partir del análisis exploratorio, la ingeniería de variables y la modelación predictiva, se identificaron **insights clave sobre qué factores influyen en la probabilidad de que un producto listado en MercadoLibre sea vendido**. Estos hallazgos pueden ser aplicados directamente por áreas de **Marketing, Category Management, Estrategia Comercial** y **Onboarding de Vendedores**, por ejemplo:

- **Estado del producto (`is_new`):**
Los productos nuevos tienden a tener mayor probabilidad de venta, probablemente por su mejor condición o novedad.
- **Engagement previo (`tags_dragged_bids_and_visits`):**
Clics, visitas y señales tempranas de interés son predictoras de ventas futuras. Acciones de marketing que generen interacción temprana pueden aumentar conversiones.
- **Disponibilidad de inventario (`available_quantity`):**
Una mayor cantidad disponible mejora la visibilidad y la probabilidad de conversión. Es recomendable evitar publicaciones con inventario muy limitado.
- **Calidad o experiencia del vendedor (`seller_loyalty_cat`):**
Los vendedores con mejores niveles de fidelidad tienden a lograr más ventas, posiblemente por su reputación, experiencia o buenas prácticas.



Caso de Uso 1: Campañas exclusivas para productos nuevos

Objetivo:

Aumentar la visibilidad y conversión de productos nuevos con alto potencial de venta.

Acción de Marketing:

- Crear campañas (e.g. "Novedades del mes") que solo incluyan productos marcados como `is_new = True`
- Utilizar etiquetas visuales tipo "¡Nuevo!" o "Recién publicado" en banners, newsletters o categorías

- Priorizar estos productos en la landing page de inicio durante los primeros días de publicación

Justificación:

Los productos nuevos muestran una mayor tasa de conversión. Resaltarlos impulsa ventas y renueva el catálogo percibido por los usuarios.

**Caso de Uso 2: Activación temprana con campañas de engagement****Objetivo:**

Fomentar señales de interacción (clics, vistas, visitas) en las primeras horas o días tras la publicación de productos.

Acción de Marketing:

- Lanzar campañas publicitarias automáticas en redes o display interno durante las primeras 24–48h de publicación
- Implementar descuentos o beneficios visibles si el producto recibe visitas tempranas
- Notificar al usuario si su producto no ha recibido interacción en las primeras horas

Justificación:

Variables como `tags_dragged_bids_and_visits` son altamente predictivas. La generación de engagement temprano puede empujar al algoritmo y aumentar la probabilidad de venta.

**Caso de Uso 3: Segmentación premium por inventario y fidelidad del vendedor****Objetivo:**

Reservar espacios de marketing premium (banners, destacados, mails) para productos que cumplen con condiciones mínimas de inventario y reputación del vendedor.

Acción de Marketing:

- Crear un filtro automático para que solo entren en campañas los productos con `available_quantity >= X` y `seller_loyalty_cat ∈ {silver, gold}`
- Enviar campañas personalizadas a compradores con productos de vendedores confiables y alto stock

- Recompensar a estos vendedores con mayor exposición en fechas clave

Justificación:

Una mayor disponibilidad y un vendedor confiable mejoran la experiencia del cliente y la tasa de conversión. Focalizar esfuerzos de marketing en este grupo optimiza el retorno.

3. Estrategia de Monitoreo

En un entorno de producción, es crítico monitorear tanto los **datos de entrada** como el **desempeño del modelo**:



Monitoreo de Datos

- Detección de *data drift* usando histogramas de distribución por campo (**price**, **is_new**, etc.)
- Alertas si cambia la proporción de la clase objetivo (**sold**) o aparecen categorías desconocidas



Monitoreo del Modelo

- Métricas en tiempo real: accuracy, F1, precision y recall en lotes diarios o semanales
 - Seguimiento de predicciones incorrectas sobre productos que sí se vendieron
 - Dashboard en **Streamlit** o **Grafana** conectado a logs de predicción
-

4. Implementación Técnica del Pipeline

[Raw CSV en GCS]

↓ (trigger: Cloud Function)

[Limpieza y Feature Engineering en Cloud Run]

↓

[Processed Dataset → BigQuery]

↓ (programado)

[Entrenamiento de modelo en Vertex AI Job o AI Platform]

↓

[Registro en MLflow Model Registry] o Vertex AI Registry



[Servicio REST para predicciones en Cloud Run o Vertex AI Endpoint]

El pipeline comienza con la carga de un archivo **CSV crudo en Google Cloud Storage (GCS)**. Esta acción **dispara automáticamente una Cloud Function**, que activa un servicio desplegado en **Cloud Run** encargado de realizar la limpieza de datos y la generación de variables (feature engineering).

Una vez procesado, el dataset final se guarda en **BigQuery**, desde donde puede ser consumido tanto por dashboards analíticos como por procesos automatizados de entrenamiento. De forma programada, un job en **Vertex AI** (o AI Platform) extrae los datos, entrena un modelo de machine learning y registra su resultado en un **Model Registry** como MLflow o Vertex AI Registry.

Finalmente, el modelo registrado puede ser desplegado como un **servicio REST en Cloud Run o Vertex AI Endpoint**, permitiendo consumir predicciones en tiempo real desde aplicaciones internas o sistemas comerciales.



Conclusión

Este proyecto permite no solo entrenar un modelo robusto con buena capacidad explicativa, sino también traducir sus resultados a **acciones concretas para equipos de negocio**, permitiendo cerrar el ciclo de valor desde los datos hacia decisiones estratégicas. Su aplicación puede mejorar conversiones, optimizar campañas, priorizar inventarios y elevar la experiencia del cliente.