

Manual de usuario VV Data Insights

1. Descripción general

Este sistema de recomendación fue desarrollado con el objetivo de generar sugerencias personalizadas de productos tanto para **clientes individuales (B2C)** como para **empresas (B2B)**. La solución implementa tres enfoques complementarios de recomendación:

- **Modelo XGBoost:** Basado en aprendizaje supervisado con variables tabulares que capturan el comportamiento histórico de compra y atributos de producto/cliente.
- **Modelo LightFM:** Utiliza aprendizaje de representaciones (factorización de matrices) para capturar relaciones implícitas entre usuarios y productos, integrando características cruzadas.
- **Modelo Híbrido:** Combina ambos enfoques mediante una ponderación ajustable (α), permitiendo balancear la influencia de cada modelo según los objetivos del negocio.

Esta versión incluye una **interfaz gráfica de validación local**, desarrollada con Dash, que permite al usuario explorar los resultados del modelo de forma dinámica. La interfaz ofrece controles para seleccionar un cliente, definir el número de recomendaciones y ajustar el valor de α para observar cómo varían los resultados.

Todos los modelos han sido entrenados y evaluados utilizando métricas estándar como **AUC, precisión, recall y F1-score**, cuyos resultados pueden consultarse en el informe técnico anexo.

2. Requisitos previos

Para ejecutar correctamente la solución de recomendación, es necesario descomprimir el archivo entregado bajo el nombre **“Algoritmo de Recomendación”**, el cual contiene todos los recursos necesarios para correr los modelos y visualizar los resultados.

Dentro de este archivo encontrarás:

- **Dos notebooks Jupyter (.ipynb)**, uno para cada enfoque:
 - *RecomendacionB2B.ipynb* para clientes B2B
 - *recomendacionB2C.ipynb* para clientes B2C
- **/modelos (Carpeta modelos):** Contiene los modelos entrenados (.pkl) y estructuras auxiliares como *features* de usuarios y productos, generadas por los notebooks:
- **/funciones (Carpeta funciones):** Scripts Python (*recomendadorB2B.py*, *recomendadorB2C.py*) que definen la lógica de recomendación utilizada por la interfaz.
- **/assets (Carpeta assets):** Archivos estáticos utilizados por Dash (por ejemplo, el logo de Corona para la interfaz).
- **app.py:** Archivo principal de la herramienta gráfica. Una vez generados los modelos (.pkl), este archivo permite lanzar la interfaz visual desarrollada con Dash.

Nota: Todo el proceso previo, de manejo de datos, entendimiento de los mismos y creación de nuevas variables está dentro del archivo comprimido “Reto Corona” enviado la entrega pasado y adjunto en esta entrega final, a partir de ese proceso se generaron las dos bases de donde se genera el modelo de recomendación.

Los archivos **RecomendacionB2B.ipynb** y **recomendacionB2C.ipynb** pueden ejecutarse desde la IDE local de preferencia. En nuestro caso, empleamos **Visual Studio Code (VS Code)**.

3. Uso paso a paso

0. Ejecutar los cuadernos para construir los modelos: Abrir y correr localmente los archivos *RecomendacionB2B.ipynb* y *recomendacionB2C.ipynb*. Estos cuadernos construyen y guardan los modelos en archivos .pkl que serán luego utilizados por la interfaz gráfica.

- **B2B** se ejecuta en menos de 1 minuto.
- **B2C** puede tardar entre **15 y 20 minutos**, según la capacidad del equipo.

Nota: Se recomienda ejecutar estos cuadernos siempre desde el equipo donde se desplegará la interfaz. Esto se debe a que los archivos .pkl pueden estar atados a versiones específicas de librerías, lo que podría causar errores al moverlos entre diferentes dispositivos o entornos. Aunque los modelos funcionen en una máquina, no hay garantía de compatibilidad inmediata en otra sin reentrenarlos.

1. Iniciar la interfaz gráfica: Una vez ejecutados los cuadernos y generados los modelos .pkl, el siguiente paso es iniciar la aplicación visual construida con **Dash**. Esta interfaz permite seleccionar parámetros y visualizar de manera dinámica las recomendaciones.

Pasos para ejecutar:

- i. Abrir el archivo App.py desde el entorno local (por ejemplo, VS Code).
- ii. Verificar que todas las dependencias estén instaladas. Si Dash no está disponible en el entorno, puede instalarse ejecutando el siguiente comando en la terminal:

➔ pip install dash

También es recomendable tener instaladas las siguientes bibliotecas, por si no están previamente configuradas:

➔ pip install pandas plotly scikit-learn lightfm

- iii. Ejecutar la aplicación corriendo el archivo App.py. Esto abrirá el servidor local en el puerto 8050 por defecto. Se puede observar un mensaje como el siguiente:

```
nataliavillegasMacbook-Pro-de-NATALIA-4: Algoritmo_Recomendacion % /usr/local/bin/python3 "/Users/nataliavillegas/Documents/NOVENO SEMESTRE (2025-11)/Corona/Algo
ritmo_Recomendacion/app.py"
/Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/site-packages/lightfm/_lightfm_fast.py:9: UserWarning: LightFM was compiled without OpenMP support. Only a single thread will be used.
warnings.warn!
Dash is running on http://127.0.0.1:8050/
* Serving Flask app 'app'
* Debug mode: on
```

Ilustración 1 Mensaje al correr el archivo App.py

Nota: Al correr la aplicación, es posible que aparezca el siguiente mensaje en la consola “*UserWarning: LightFM was compiled without OpenMP support. Only a single thread will be used*”. Este *warning* es solo una advertencia de que la librería LightFM no está utilizando múltiples hilos (*threads*) para procesamiento paralelo debido a una configuración de compilación. No afecta el funcionamiento, precisión ni el despliegue de la herramienta, y puede ser ignorado.

- iv. Ingresar a ese enlace desde cualquier navegador para comenzar a usar la interfaz.

Nota: Asegúrese de tener los archivos .pkl, .csv y .json generados en la misma carpeta raíz o en las rutas especificadas dentro de App.py. (Si previamente se descomprimió la carpeta y se ejecutaron los cuadernos de Jupyter, no debería tener problemas con las rutas de los archivos)

2. **Interacción con la herramienta gráfica:** Una vez ejecutado el archivo App.py, se desplegará una **interfaz web interactiva** en el navegador (por defecto en <http://127.0.0.1:8050/>), la cual permite visualizar y controlar el sistema de recomendación híbrido.

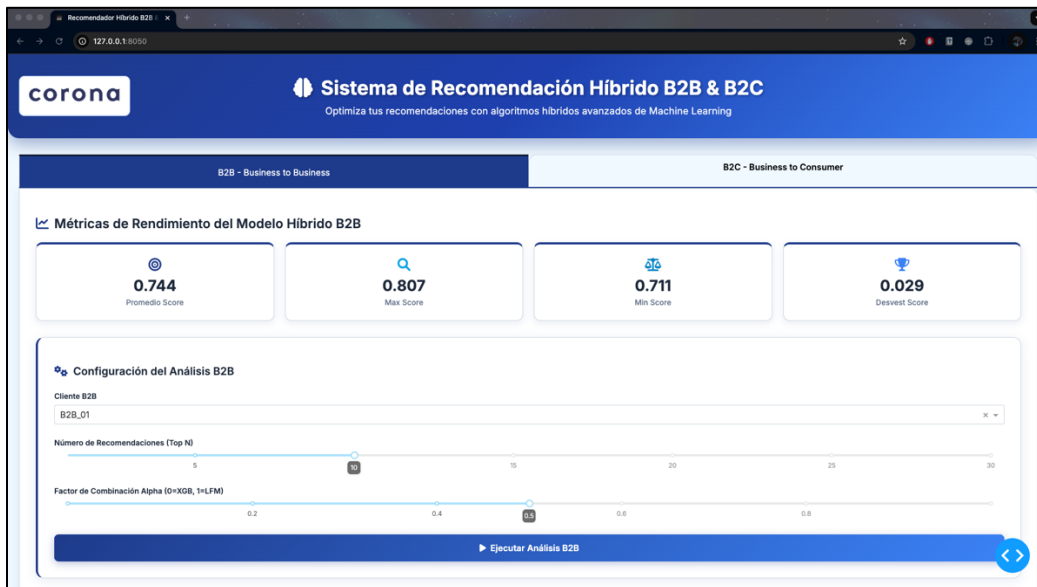


Ilustración 2 Primera interacción con la herramienta

Elementos clave de la interfaz:

- **Selector de pestañas (B2B / B2C):** En la parte superior, puede alternar entre los dos segmentos de análisis: ***Business to Business (B2B)*** o ***Business to Consumer (B2C)***. Esto ajusta la configuración y las métricas al tipo de cliente seleccionado.
- **Métricas de Rendimiento del Modelo Híbrido:** Cuatro tarjetas muestran los valores calculados previamente del desempeño del modelo híbrido:
 - Promedio *Score*
 - Máximo *Score*
 - Mínimo *Score*
 - Desviación estándar del *Score*

Estas métricas brindan una visión general de la calidad y dispersión de las recomendaciones generadas.

- **Configuración del Análisis:** Un panel que permite al usuario personalizar la recomendación:
 - **Cliente:** Selección del ID del cliente (B2B o B2C según el segmento activo).
 - **Número de Recomendaciones (Top N):** Slider para definir cuántos productos mostrar (entre 1 y 30).
 - **Factor de Combinación Alpha (0 = XGB, 1 = LFM):** Control deslizante que ajusta el peso relativo de cada modelo en la combinación híbrida.
- **Botón de Ejecución:** Al pulsar "Ejecutar Análisis", se actualizan automáticamente todos los gráficos y la tabla de resultados, reflejando las recomendaciones personalizadas para ese cliente y configuración.

Nota: Al ejecutar el script por primera vez, es posible que aparezca un mensaje como: "*LightFM was compiled without OpenMP support. Only a single thread will be used*". Esto es solo una advertencia sobre paralelización y no afecta el funcionamiento, rendimiento o despliegue correcto de la herramienta.

4. Resultados presentados

Una vez ejecutado el análisis (haciendo clic en el botón correspondiente), se despliegan automáticamente los siguientes resultados en la interfaz:

a. Gráfico 1: Comparación de Scores por Modelo

Un gráfico de barras agrupado por producto recomendado que permite comparar, para cada uno:

- Score del modelo LightFM (azul oscuro).
- Score del modelo XGBoost (azul claro).
- Score Híbrido (azul medio).

Este gráfico permite entender cómo se comporta cada modelo de manera individual, identificando cuál de ellos está aportando más peso en la recomendación final de cada producto. Además, facilita el análisis del impacto del parámetro α sobre la recomendación híbrida. Cabe resaltar que los productos se presentan ordenados de izquierda a derecha, donde el producto ubicado más a la izquierda corresponde al más recomendado (mayor score), y a medida que se avanza hacia la derecha, la probabilidad estimada de intención de compra disminuye.

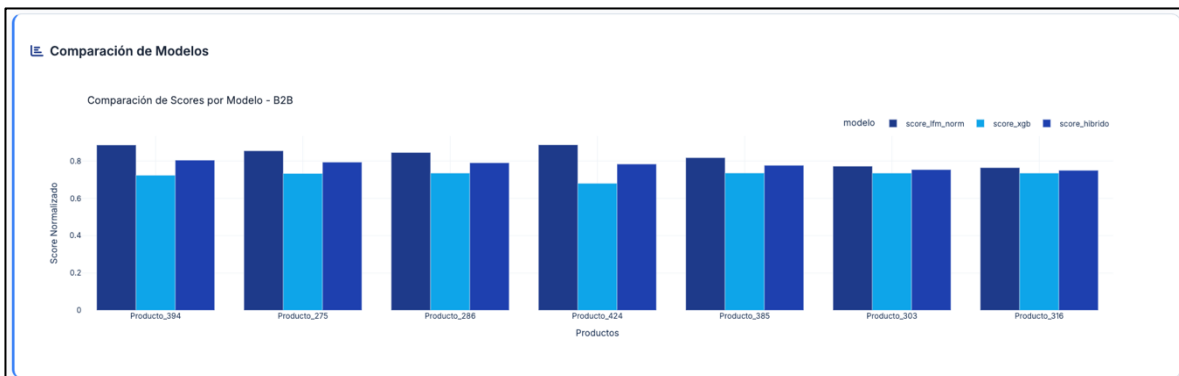


Ilustración 3 Gráfica de barras comparativas de los modelos

b. Gráfico 2: Ranking de Productos por Score Híbrido

Un gráfico de barras horizontal que presenta los productos recomendados ordenados según su score híbrido.

- Muestra los productos de menor a mayor score híbrido.
- Cada barra representa un producto, y su longitud refleja el nivel de recomendación.
- Utiliza una escala de color para reforzar visualmente la intensidad del score.

Este gráfico permite visualizar rápidamente cuáles son los productos más recomendados para el cliente seleccionado, destacando aquellos con mayor probabilidad de intención de compra según el modelo híbrido. Su diseño facilita comparar las diferencias de puntuación entre productos, ayudando a priorizar decisiones comerciales de manera efectiva.

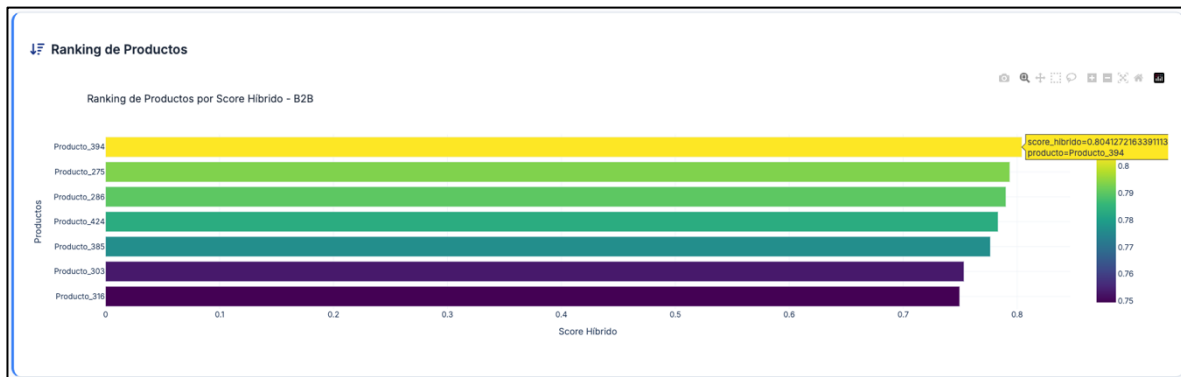


Ilustración 4 Ranking de los productos recomendados

c. Gráfico 3: Valor Esperado vs Alineación Estratégica

Un gráfico de dispersión que combina criterios estratégicos y económicos con el resultado del modelo.

- Eje X: alineación estratégica del producto (estrategia comercial o de portafolio)
- Eje Y: valor económico esperado del producto para el cliente.
- Color del punto: score híbrido (más oscuro = mayor score).
- Tamaño del punto: magnitud del score híbrido (más grande = más relevante).

Este gráfico permite identificar productos que no solo tienen un alto score desde el modelo, sino que también están alineados con la estrategia de la empresa y representan alto valor económico. Es especialmente útil para decisiones de negocio más sofisticadas, ya que combina el output del algoritmo con variables clave de negocio.

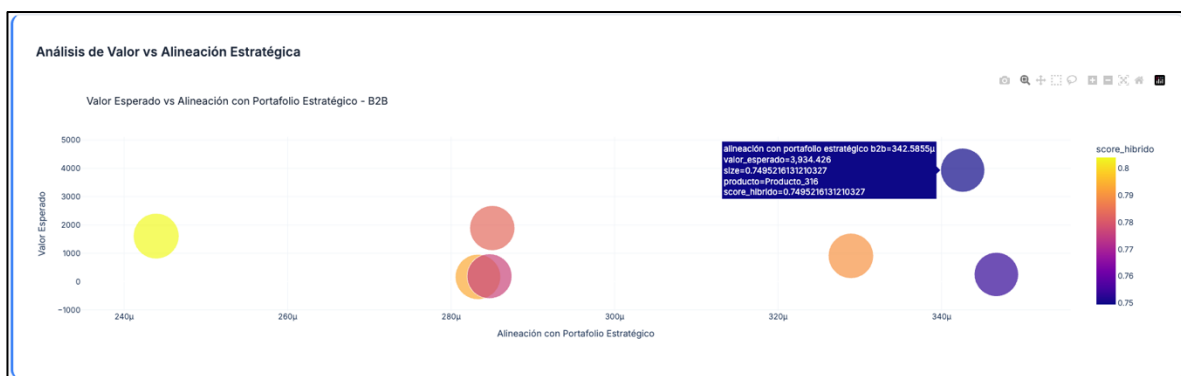


Ilustración 5 Gráfica comparativa entre valor esperado y alineación estratégica

d. Tabla de Recomendaciones Personalizadas

Se muestra una tabla ordenada de mayor a menor según el **score híbrido** de recomendación para los productos seleccionados. Esta tabla incluye:

- **Producto:** Nombre o ID del producto.

- **Score LFM Norm:** Puntaje generado por el modelo LightFM (normalizado entre 0 y 1).
- **Score XGB:** Puntaje generado por el modelo XGBoost.
- **Score Híbrido:** Combinación lineal entre los dos scores anteriores, ponderado según el parámetro α .
- **Alineación Estratégica:** Grado de afinidad del producto con el portafolio estratégico.
- **Valor Esperado:** Valor económico estimado de la compra del producto.

Producto	Score LFM Norm	Score XGB	Score Híbrido	Alineación Estratégica	Valor Esperado
Producto_394	0.885	0.723	0.804	0.000244	1605.88
Producto_275	0.854	0.733	0.794	0.000283	164.93
Producto_286	0.845	0.735	0.790	0.000329	907.78
Producto_424	0.886	0.680	0.783	0.000285	1890.00
Producto_385	0.817	0.735	0.776	0.000285	188.58
Producto_303	0.772	0.735	0.753	0.000347	252.34
Producto_316	0.764	0.735	0.750	0.000343	3934.43

Ilustración 6 Tabla resumen de resultados

5. Consideraciones adicionales

- **Interactividad dinámica:** El sistema es completamente reactivo; cada vez que se cambia el ID del cliente, el número de recomendaciones o el parámetro α (*alpha*), se recalculan y actualizan automáticamente los resultados en tiempo real.
- **Ejecución manual del análisis:** Para generar las recomendaciones y visualizar los gráficos correspondientes, es necesario presionar el botón “Ejecutar Análisis” después de configurar los parámetros. Esto asegura que los cálculos se realicen bajo las condiciones seleccionadas.
- **Transición entre pestañas:** Al cambiar entre las pestañas B2B y B2C, puede haber una pequeña demora (unos pocos segundos), especialmente al cargar la vista B2C. Esto se debe al tamaño considerable de la base de datos B2C, lo cual puede afectar ligeramente el tiempo de renderizado del tablero.
- **Métricas precomputadas:** Las métricas de evaluación del modelo (Precisión, Recall, F1, AUC, etc.) no se recalculan en tiempo real, sino que fueron evaluadas previamente y se cargan desde estructuras predefinidas.
- **Escalabilidad del sistema:** El sistema está preparado para recibir nuevas actualizaciones o bases de datos adicionales sin requerir cambios estructurales importantes. Esto permite continuar el desarrollo y escalar la solución a futuro.

- Archivos complementarios: Dentro del documento comprimido entregado se encuentran archivos adicionales que no son utilizados directamente por esta interfaz. Estos corresponden a fases anteriores del proyecto
- **Próximo paso lógico:** Con la interfaz final validada, el siguiente paso natural es su **despliegue en la nube**, permitiendo así una disponibilidad continua, integración con otros sistemas internos, mayor capacidad de cómputo, y acceso multiusuario. Esto permitirá que áreas como comercial o mercadeo también puedan aprovechar la herramienta de forma sencilla y segura.