

Análisis de Ventas y Segmentación de Clientes en un E-commerce

Autor: Italo Buonanno.

Mentora: Alana Olivieri.

Resumen

El objetivo de este estudio es analizar el comportamiento de los usuarios en un entorno de e-commerce, con el fin de segmentarlos en distintas categorías para su aplicación en estudios de mercado. Para ello, se realizó un análisis descriptivo de las ventas, se aplicó la segmentación Recency, Frequency, Monetary (RFM) y se empleó un modelo de Random Forest para predecir la categoría de los clientes.

El análisis descriptivo permitió identificar diferencias entre los productos más vendidos en términos brutos y netos, destacando el impacto de las devoluciones en la valoración de los productos. Se observó una reducción progresiva en el número de devoluciones y un aumento en las ventas a lo largo del tiempo. Asimismo, se evidenció que los clientes con mayor gasto tienden a registrar menos devoluciones. La segmentación RFM mostró que la mayoría de los usuarios pertenecen a una categoría mixta, reflejando la diversidad de comportamientos de compra. El modelo de Random Forest obtuvo un alto nivel de precisión en la predicción de las categorías de clientes según los criterios definidos.

Estos resultados permiten profundizar en el conocimiento del comportamiento del cliente y constituyen una base sólida para el diseño de estrategias empresariales fundamentadas en datos.

Introducción

El presente estudio tiene como objetivo general:

Analizar el comportamiento de los usuarios en un entorno de E-commerce para segmentarlos en distintas categorías relevantes para un estudio de mercado.

Y como objetivos específicos:

1. Analizar la cantidad total de productos vendidos.
2. Examinar las cantidades vendidas por mes.
3. Analizar el comportamiento de compra de los usuarios.

4. Agrupar a los clientes según su comportamiento de compra utilizando el modelo RFM.
5. Aplicar el modelo de Random Forest para predecir la categoría de los clientes.

Metodología

En primer lugar, se realizó un análisis descriptivo de los datos. Posteriormente, se aplicó la segmentación RFM para categorizar a los usuarios según su comportamiento de compra. Finalmente, se implementó un modelo de Machine Learning, utilizando Random Forest, con el objetivo de predecir la categoría de los clientes a partir de las variables disponibles.

Resultados

Se observó que el producto más vendido en términos brutos (sin descontar devoluciones) no coincide con el producto más vendido de forma neta (descontando las devoluciones). Además, los primeros meses registraron un número elevado de devoluciones, cifra que fue disminuyendo progresivamente al mismo tiempo que aumentaban las ventas. Se evidenció también que los clientes con mayor gasto presentan un bajo nivel de devoluciones.

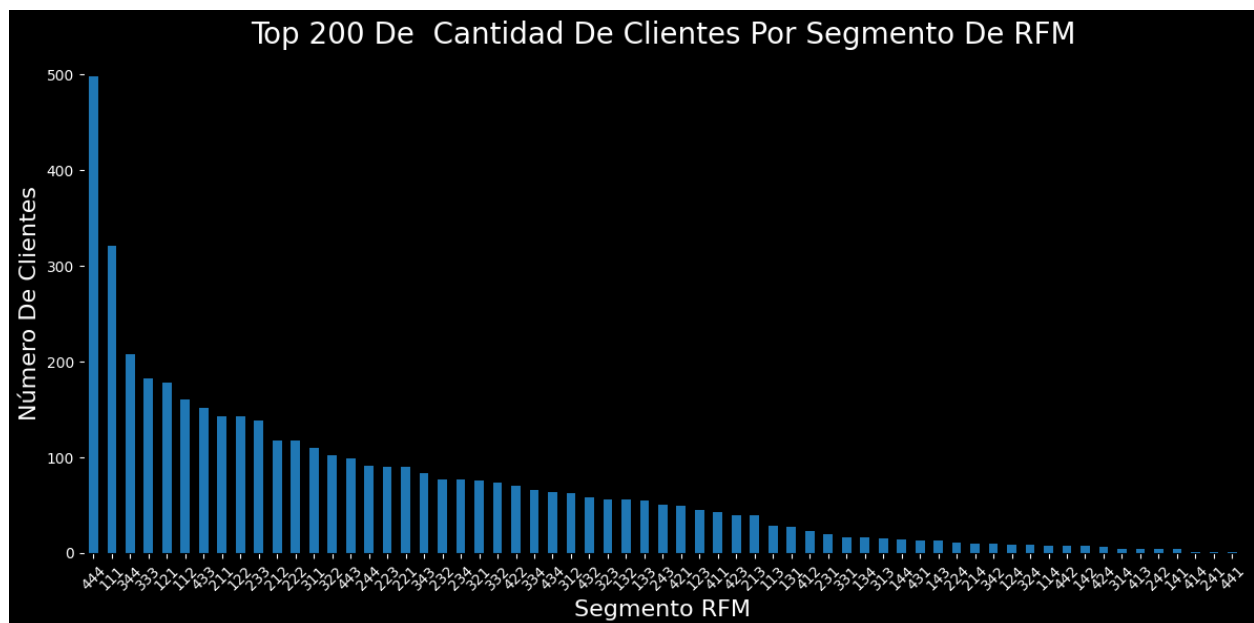


Figura 1: Top 200 De Cantidad De Clientes Por Segmento De RFM.

Como se observó en la Figura 1, la segmentación RFM mostró que la mayoría de los usuarios pertenecen a una categoría mixta.

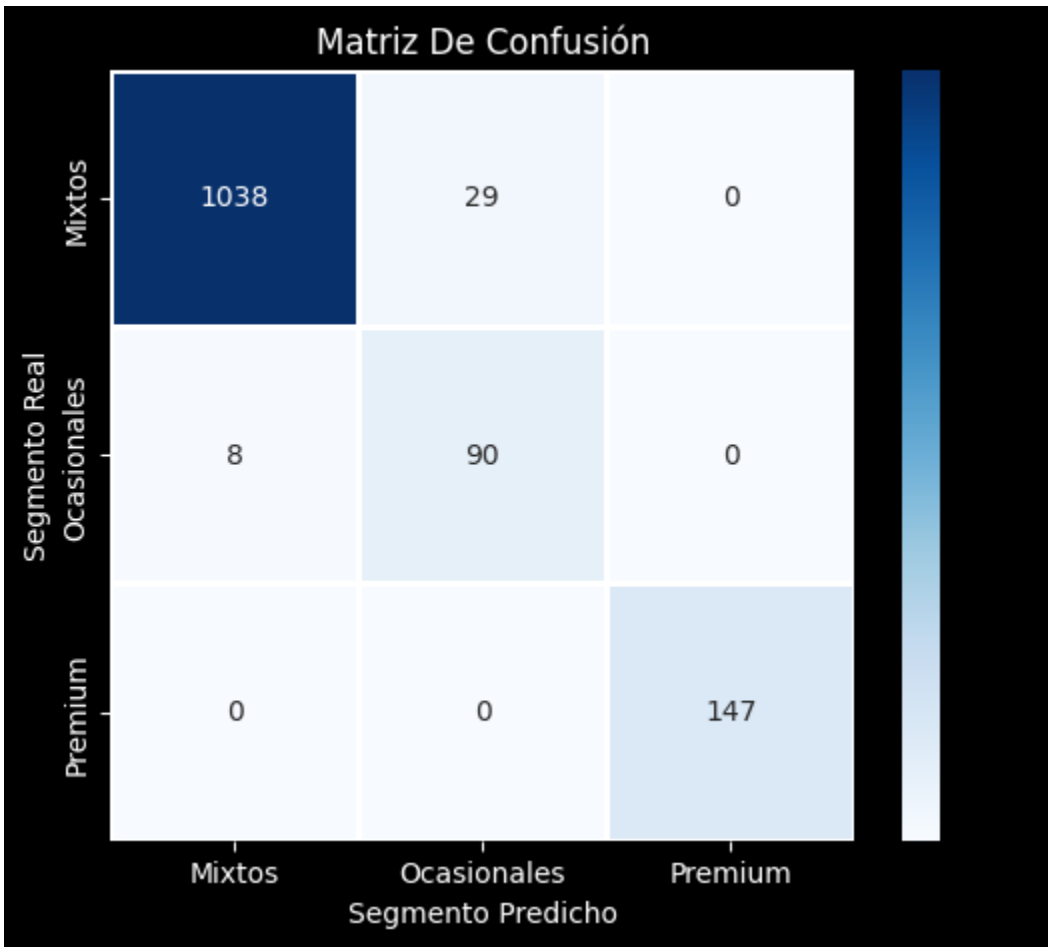


Figura 2: Matriz de Confusión.

Finalmente, el modelo de Random Forest alcanzó un elevado nivel de aciertos en la predicción de la categoría de los clientes, según los criterios encontrados, siendo estos evidente en la Figura 2.

Discusión

El análisis de las ventas netas resulta útil para evaluar la rentabilidad de un producto, mientras que el estudio de las ventas brutas permite identificar posibles problemas asociados a productos específicos. La segmentación por RFM es especialmente valiosa para implementar estrategias de marketing más personalizadas, especialmente en mercados donde existe una gran variabilidad en el comportamiento de los clientes, permitiendo así obtener resultados más efectivos.

Conclusión

Este estudio ha demostrado la utilidad de combinar el análisis descriptivo, la segmentación RFM y los modelos predictivos para comprender el comportamiento de los clientes en e-commerce. Considerar las devoluciones es clave para evaluar la rentabilidad real de los productos, y la segmentación permite personalizar estrategias comerciales. Se recomienda continuar explorando técnicas avanzadas de segmentación y predicción para mejorar la toma de decisiones y maximizar resultados empresariales.

Referencias

Conecta Software. (s.f.). *RFM: Recency, Frequency, Monetary*. Recuperado de <https://www.conectasoftware.com/magazine/glosario/rfm/>

IBM. (s.f.). *Random Forest*. Recuperado de <https://www.ibm.com/think/topics/random-forest>

UCI Machine Learning Repository. (s.f.). *Online Retail Data Set*. Recuperado de <https://archive.ics.uci.edu/dataset/352/online+retail>