



Estimación del Customer Lifetime Value en un Contexto No Contractual

Jhon Jarrys Vásquez Zapata

Informe de práctica para optar el título de Ingeniero de Sistemas

Asesor

Daniela Serna Buitrago

Especialista en Analítica y Ciencia de Datos

Universidad de Antioquia

Facultad de Ingeniería

Pregrado

Medellín

2024

Cita

(Vásquez Zapata, 2024)

Referencia

Vásquez Zapata, J. J. (2024). *Estimación del Customer Lifetime Value en un Contexto No Contractual* [Práctica Empresarial]. Universidad de Antioquia, Medellín

Estilo APA 7 (2020)



Centro de Documentación Ingeniería (CENDOI)

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano/Director: Julio César Saldarriaga Molina.

Jefe departamento: Diego José Luis Botía Valderrama.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Dedicatoria

A mí familia, en especial mi madre y abuela, quienes han sido fundamentales en mi crecimiento personal y profesional.

Agradecimientos

Quisiera agradecer a Hermeco por abrirme las puertas para realizar estas prácticas, las cuales me han permitido adquirir un conocimiento invaluable. También a Daniela Serna por su excepcional asesoramiento y orientación durante todo este proceso.

Tabla de contenido

Resumen	8
Abstract	9
Introducción	10
1 Objetivos	11
1.1 Objetivo general	11
1.2 Objetivos específicos	11
2 Marco teórico	12
2.1 Customer Lifetime Value	12
2.2 Modelos BTYD	13
2.2.1 Modelo RFM	16
2.2.2 Modelo Gamma-Gamma	16
3 Metodología	17
3.1 Recolección y manipulación de datos	17
3.2 Entrenamiento de modelos	18
4 Resultados	20
5 Análisis	26
6 Conclusiones	27
Referencias	28

Lista de tablas

Tabla 1 Mejores valores para el coeficiente de penalización y error cuadrático medio	20
---	----

Lista de figuras

Figura 1	Clasificación de la relación empresa-cliente	13
Figura 2	Diversos historiales de compra	14
Figura 3	Datos convertidos a estructura RFM	18
Figura 4	División del conjunto de datos	19
Figura 5	Frecuencia de transacciones repetidas para el modelo BG/NBD	21
Figura 6	Frecuencia de transacciones repetidas para el modelo MBG/NBD	21
Figura 7	Compras reales en el período de reserva frente a compras predichas (BG/NBD)	22
Figura 8	Compras reales en el período de reserva frente a compras predichas (MBG/NBD)	22
Figura 9	Matriz para analizar las futuras compras respecto a la frecuencia y recencia	23
Figura 10	Correlación entre las variables frecuencia y valor monetario	24
Figura 11	Función para estimar el CLV en un año	24
Figura 12	Resultado final de la estimación del CLV	25

Siglas, acrónimos y abreviaturas

CLV	Customer Lifetime Value
BG/NBD	Beta-Geometric/Negative-Binomial-Distribution
MBG/NBD	Modified Beta-Geometric/Negative-Binomial-Distribution
BTYU	Buy 'Till You Die

Resumen

Identificar a los clientes más valiosos es esencial para cualquier empresa, ya que permite desarrollar estrategias comerciales más efectivas y fortalecer la fidelización. Una métrica fundamental para lograrlo es el Valor de Vida del Cliente (Customer Lifetime Value). En contextos no contractuales, estimar esta métrica puede resultar complejo debido a la incertidumbre inherente. Por lo cual una de las opciones es utilizar modelos probabilísticos para abordar esta cuestión.

En este documento se presenta la aplicación de los modelos BG/NBD, MBG/NBD y Gamma-Gamma en datos transaccionales de una marca de ropa para niños y adolescentes. El modelo implementado arrojó resultados satisfactorios para la estimación del Customer Lifetime Value y el análisis del comportamiento de los clientes de la empresa. Este enfoque analítico no solo permite estimar la cantidad de compras que realizará un cliente en determinado momento o el gasto promedio por transacción, sino que también ofrece una visión más profunda del comportamiento del cliente y sus preferencias, lo que facilita la toma de decisiones estratégicas.

Palabras clave: valor de vida del cliente, modelo RFM, mercadotecnia, modelos probabilísticos.

Abstract

Identifying the most valuable customers is essential for any company, as it allows for the development of more effective business strategies and strengthens customer loyalty. A fundamental metric for achieving this is Customer Lifetime Value (CLV). In non-contractual contexts, estimating this metric can be complex due to inherent uncertainty, hence one of the options is to use probabilistic models to address this issue.

This document presents the application of BG/NBD, MBG/NBD, and Gamma-Gamma models on transactional data from a children's and teenagers' clothing brand. The implemented model yielded satisfactory results for estimating Customer Lifetime Value and analyzing the behavior of the company's customers. This analytical approach not only allows for estimating the amount of purchases a customer will make at a given time or the average expenditure per transaction but also provides a deeper insight into customer behavior and preferences, facilitating strategic decision-making.

Keywords: customer lifetime value, RFM model, marketing, probabilistic models.

Introducción

Conservar relaciones sólidas y duraderas con los clientes es fundamental para cualquier empresa. La gestión efectiva de estas relaciones se ha vuelto esencial en la estrategia comercial (Hu et al., 2013). Por tanto, es importante conocer el valor de cada cliente y cuanto beneficio a futuro generará. Medir y gestionar estos valores ofrece ventajas competitivas a una empresa (Zhang et al., 2010), al permitirle fortalecer sus vínculos con los clientes y tomar decisiones estratégicas que impulsen el crecimiento y la rentabilidad a largo plazo.

Por esta razón, como parte de su estrategia comercial, la compañía Hermeco desea implementar nuevas estrategias en el análisis de datos para conocer cuáles son sus clientes más valiosos. Este nuevo proceso implica identificar aquellos clientes que no solo generan ingresos significativos en el presente, sino también aquellos que tienen el potencial de generar beneficios sostenidos en el futuro. Al comprender mejor el valor de cada cliente, la empresa puede dirigir sus recursos y esfuerzos hacia la retención y el fortalecimiento de las relaciones con estos clientes clave, además de otras estrategias empresariales.

Una métrica importante en este contexto es el Valor de Vida del Cliente (Customer Lifetime Value), la cual representa el valor monetario que un cliente aporta a una empresa durante toda su relación comercial. Estimar esta métrica proporciona una comprensión más profunda del comportamiento futuro de cada cliente y las ganancias que generará, lo que permite tomar decisiones estratégicas más informadas y acertadas.

Por consiguiente, en este proyecto se emplean modelos probabilísticos como BG/NBD y MBG/NBD para proyectar la cantidad de compras que un cliente realizará en el futuro. Además, se implementa un modelo Gamma-Gamma para calcular la cantidad promedio que gasta un cliente por transacción. Al combinar estos dos modelos, se logra estimar el Customer Lifetime Value con un alta precisión.

1 Objetivos

1.1 Objetivo general

Estimar el valor del CLV de cada cliente de la compañía, con el fin de optimizar las estrategias de marketing y retención de clientes.

1.2 Objetivos específicos

- Analizar el comportamiento histórico de compra de los clientes para identificar patrones y tendencias que influyan en el valor del CLV.
- Crear modelos predictivos utilizando datos transaccionales para estimar el CLV de cada cliente de la empresa.
- Identificar ventajas y desventajas en la elección de un modelo específico para la predicción del Customer Lifetime Value, teniendo en cuenta criterios como complejidad del modelo, calidad de datos y su capacidad de generalización.
- Limpiar y tratar los datos provenientes de las diferentes bases de datos de la compañía

2 Marco teórico

2.1 Customer Lifetime Value

El Customer Lifetime Value (CLV) es una métrica utilizada para predecir el comportamiento futuro del cliente y dirigirse a aquellos que son más rentables (Mzoughia y Limam, 2014). Esta métrica se define como el valor presente de todos los beneficios futuros que un cliente generará durante su relación con una empresa (Nikkhahan, Badrabadi, & Tarokh, 2011). El Customer Lifetime Value no solo proporciona una mejor perspectiva del valor de cada cliente, sino que también sirve como brújula estratégica para la toma de decisiones empresariales.

Existen diversos enfoques de modelado del CLV. Gupta et al. (2006) describieron seis enfoques: modelos RFM, modelos de probabilidad, modelos econométricos, modelos de persistencia, modelos de ciencias de la computación y modelos de difusión y crecimiento. Los modelos RFM están basados en la Recencia, Frecuencia y Valor Monetario; los modelos de probabilidad en los modelos Buy 'Till You Die, que abarcan los modelos Pareto/NBD y BG/NBD; los modelos econométricos buscan describir el comportamiento de los consumidores en función de una serie de covariables; los modelos de persistencia se centran en modelar el comportamiento de retención, adquisición y venta cruzada; la amplia literatura de ciencias de la computación ha generado muchos enfoques centrados en la capacidad predictiva, estos incluyen modelos como las redes neuronales o árboles de decisión; los modelos de difusión y crecimiento están basados en "Customer Equity" (CE).

Ahora bien, para realizar una estimación correcta del CLV se debe considerar el tipo de relación que tienen los clientes con la compañía. En una relación contractual, los clientes informan a la compañía cuando deciden finalizar dicha relación. Sin embargo, en una relación no contractual, la compañía debe inferir si un cliente aún está activo (Gupta et al., 2006). Por ejemplo, una suscripción mensual a Netflix se considera una relación contractual, ya que la cancelación del servicio implícitamente indica la inactividad del cliente. En contraste, en una tienda de ropa, determinar si un cliente sigue activo es más complejo, ya que es incierto si el cliente volverá a realizar una compra después de su última transacción.

De igual manera, es importante tener en cuenta la tipología de las transacciones de productos. Las transacciones continuas son aquellas que pueden ocurrir en cualquier momento; las transacciones discretas, por el contrario, solo pueden ocurrir en un punto específico del tiempo (Martins, 2021). La clasificación de las relaciones entre clientes-compañía y la tipología de las transacciones se observan en la **Figura 1**.

Figura 1.

Clasificación de la relación empresa-cliente

Oportunidad de las transacciones	Continua	Compras en supermercados Alojamiento en hoteles	Tarjeta de Credito Plan celular
	Discreta	Asistencia a un evento Prescripción médica	Poliza de seguros Suscripción a revistas
		No contractuales	Contractuales

Tipo de relación con los clientes

Nota. Adaptada de (Fader y Hardie, 2009).

2.2 Modelos BTYD

Un enfoque probabilístico aborda el problema de modelado con la premisa de que el comportamiento observado de los clientes es un proceso estocástico subyacente. Esto implica que siempre habrá incertidumbre en cuanto a cuándo un cliente realizará su próxima compra. Por ejemplo, si un cliente ha realizado numerosas compras en el pasado, nada garantiza que vuelva a comprar en el futuro (Fader y Hardie, 2009).

El conjunto de modelos probabilísticos BTYD (Buy ‘till you die) son los más representativos para predecir el comportamiento de los clientes en entornos no contractuales (Kim

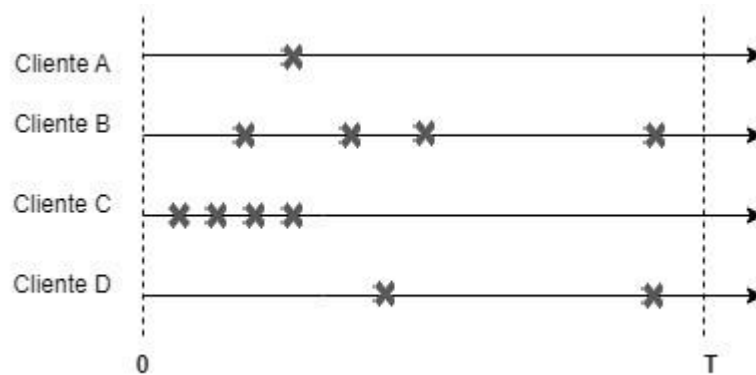
et al., 2022). Un objetivo de esta clase de modelos es, dado su historial de compras, predecir qué clientes tienen más probabilidad de estar “vivos” (Fader y Hardie, 2009). Si un cliente tiene una alta probabilidad de estar “vivo”, entonces la empresa asumirá que es probable que realice compras en el futuro.

En un contexto no contractual, se presentan diversos comportamientos por parte de los clientes, lo que resulta en una gran variedad de historiales de compra. Esto se puede apreciar en la **Figura 2**, donde T representa el momento en el tiempo en el que se desea determinar la probabilidad de que un cliente esté “vivo” y cada “x” sobre las líneas de tiempo es el momento en que cada cliente realizó una compra. En este ejemplo, a priori no es posible saber si un cliente está vivo en ese instante de tiempo, pero es posible realizar un primer análisis del comportamiento de los clientes:

- Los clientes B y D realizaron la última compra en el mismo tiempo, por lo que pueden tener una probabilidad similar de estar activos en el instante T . Sin embargo, dado que D ha realizado menos compras en el pasado, se podría esperar un menor número de compras de este cliente en el futuro en comparación con B.
- En el intervalo de tiempo $(0, T]$ los clientes B y C realizaron la misma cantidad de compras. Sin embargo, las transacciones de C son más antiguas, lo que sugiere que tiene una probabilidad menor de estar activo en comparación con B (Fader y Hardie, 2009).

Figura 2.

Diversos historiales de compra



Nota. Adaptada de (Fader y Hardie, 2009).

Uno de los primeros modelos BTYD propuestos fue el modelo Pareto/NBD, el cuál fue desarrollado por Schmittlein, Morrison y Colombo (1987). Este modelo está basado en las siguientes suposiciones:

- Durante un tiempo no observado el cliente estará “vivo” para la empresa, luego se vuelve permanentemente inactivo.
- Cuando el cliente esté “vivo”, el número de compras que hace el cliente sigue un proceso de Poisson.
- La heterogeneidad en la tasa de transacciones entre los clientes se distribuye según una distribución gamma.
- El “tiempo de vida” no observado de cada cliente sigue una distribución exponencial.
- La heterogeneidad en las tasas de abandono entre los clientes sigue una distribución gamma.
- Tanto la tasa de transacciones como la de abandono varían de forma independiente entre los clientes (Gupta et al., 2006).

Sin embargo, este modelo ha tenido dificultades para su implementación, en especial por su complejidad y demanda computacional requerida para estimar sus parámetros (Fader et al., 2005). Es por esto por lo que Fader et al., (2005) propusieron una alternativa llamada modelo BG/NBD. Este modelo se diferencia del anterior en como hace la suposición de que un cliente se vuelve inactivo. Como se especificó anteriormente, el modelo Pareto/NBD supone que la deserción de un cliente puede ocurrir durante un tiempo no observado, en cambio el modelo BG/NBD asume que la deserción ocurre inmediatamente después de la compra.

No obstante, el modelo BG/NBD tiene una limitación: solo considera a los clientes que tienen compras repetidas, es decir, aquellos para los cuales su frecuencia no es igual a 0. Por consiguiente, Batislam et al, (2007) modificaron dicho modelo introduciendo una condición adicional de abandono, para de esta manera poder trabajar con los clientes sin compras repetidas; este es el modelo MBG/NBD.

2.2.1 Modelo RFM

Muchos modelos probabilísticos, incluidos los modelos BTYD, parten de datos estructurados con variables de Recencia, Frecuencia y Valor Monetario (Montero Montero y Suárez Villamar, 2019). A pesar de que hay métodos para estimar el CLV usando un modelo RFM, en este contexto solo nos interesa las tres características que dan nombre a este modelo y que son usadas por los modelos BTYD para estimar sus parámetros.

Por tal razón, es importante comprender a detalle qué representan estas tres características: La Recencia (R) se refiere al intervalo de tiempo entre la primera y última compra. Por otro lado, la Frecuencia (F) se refiere al número de transacciones en un período determinado (ya sea anual, trimestral, mensual, etc). El Valor Monetario (M) representa el monto de dinero gastado en un periodo específico; a mayor gasto, mayor será el valor de M.

2.2.2 Modelo Gamma-Gamma

Para calcular el CLV no basta con conocer el número de transacciones que realizará un cliente en el futuro; también es necesario saber cuánto dinero gastará dicho cliente. Para abordar esta necesidad, se ha desarrollado el modelo Gamma-Gamma. Este modelo se utiliza para estimar el gasto medio por transacción de cada cliente (Sobreiro et al, 2019). Sin embargo, es importante tener en cuenta que este modelo hace una suposición para poder calcular de forma correcta el valor monetario: no debe haber correlación entre el Valor Monetario y la Frecuencia.

3 Metodología

Debido a su amplia adopción y eficacia en la resolución de problemas similares, la librería Lifetimes (Davidson-Pilon et al., 2019) fue seleccionada como herramienta clave para la implementación de los modelos BTYD. Esta librería ofrece funciones para el tratamiento de datos, ajuste de parámetros de los modelos y análisis del comportamiento de los clientes. Otras librerías de Python usadas son Pandas, Numpy y Matplotlib.

3.1 Recolección y manipulación de datos

Se extrajo el historial de compras de clientes desde el primero de enero de 2017 hasta el 11 de noviembre de 2023, centrándose únicamente en aquellos clientes que realizaron su primera compra a partir de la mencionada fecha inicial. Esta medida se adoptó con el objetivo de asegurar abarcar el historial completo de cada cliente. Se excluyeron los clientes mayoristas debido a la naturaleza de sus transacciones, que suelen tener movimientos en grandes volúmenes, lo cual puede sesgar los modelos.

Se realizó un análisis exploratorio de los datos para comprender mejor las distribuciones y características de los datos. Durante este proceso, se identificaron y trataron problemas típicos en la manipulación de datos, como por ejemplo los valores nulos y erróneos, contribuyendo así a mejorar la calidad y la integridad de los datos.

Para implementar este tipo de modelos, es fundamental transformar los datos transaccionales a una estructura que incorpore variables de recencia, frecuencia y valor monetario (RFM). La librería Lifetimes proporciona herramientas para realizar esta conversión de datos. La estructura final de los datos después de su conversión se muestra en la **Figura 3**, donde se puede observar el identificador del cliente (SAP), la frecuencia (frequency), la recencia (recency), el valor monetario (monetary_value) y el valor T, el cuál es la duración entre la primera compra del cliente y el final del periodo de estudio.

Figura 3.

Datos convertidos a estructura RFM

	frequency	recency	T	monetary_value
SAP				
0000100006	2.00	1320.00	1877.00	148723.00
0000100017	0.00	0.00	998.00	0.00
0000100021	1.00	1034.00	2449.00	61212.00
0000100038	0.00	0.00	1062.00	0.00
0000100040	1.00	585.00	991.00	126042.00

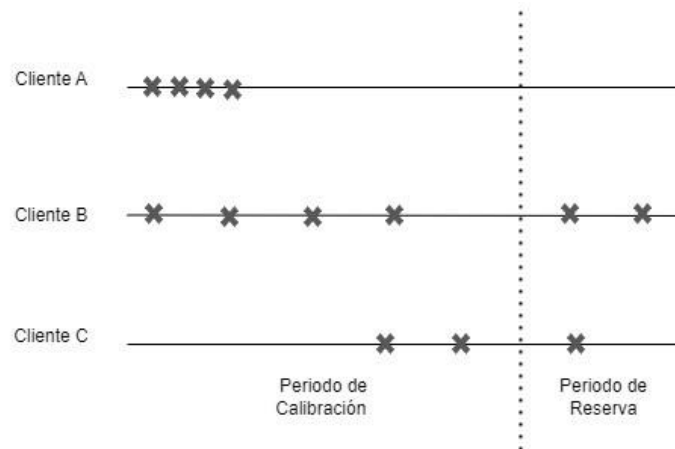
3.2 Entrenamiento de modelos

Se optó por emplear los modelos BG/NBD y MBG/NBD, los cuales gozan de un amplio reconocimiento en la literatura y han demostrado ser eficaces en la estimación del CLV y el comportamiento futuro del cliente (Batislam et al, 2007).

Para realizar el entrenamiento de los modelos, se dividió el conjunto de datos en un periodo de calibración y un periodo de reserva, para así validar las predicciones que realiza cada modelo. El periodo de calibración abarcó desde el 1 de enero de 2017 hasta el 13 de octubre de 2021, mientras que el periodo de reserva comprendió desde el 14 de octubre de 2021 hasta el 1 de noviembre de 2023. Esta división se ilustra en la **Figura 4**.

Figura 4.

División del conjunto de datos



Por otro lado, para entrenar el modelo Gamma-Gamma, se verificó la independencia entre el Valor Monetario y la Frecuencia, dado que este supuesto es fundamental para la estimación del CLV. En caso de que los datos no cumplieran con este supuesto, tocaría buscar otras alternativas para la solución del problema.

4 Resultados

Se realizó tanto el entrenamiento como la posterior evaluación de los modelos BG/NB y MBG/NBD. Estos modelos usan un hiperparámetro que introduce una penalización L2. Dado que este hiperparámetro puede variar en un rango de 0 a 1, el espacio de búsqueda es bastante amplio. Es por esto por lo que se utiliza la librería Optuna para la optimización de hiperparámetros, utilizando el Error Cuadrático Medio como métrica de evaluación. En la **Tabla 1** se pueden observar los resultados de este proceso.

Tabla 1

Mejores valores para el coeficiente de penalización y error cuadrático medio

Modelo	Mejor valor coeficiente de penalización	Error Cuadrático Medio
BG/NBD	0.0	1.96
MBG/NBD	0.0	1.92

Sin embargo, solo esta métrica no nos garantiza que los modelos están haciendo un buen trabajo. Una forma de validar la precisión de nuestros modelos en capturar el comportamiento de compra de los clientes es contrastar las frecuencias entre nuestros datos reales de calibración y los datos artificiales muestreados de las distribuciones generadas tanto por el modelo BG/NBD como por el MBG/NBD, lo cual se puede apreciar en las **Figuras 5 y 6**. Como se puede observar, ambos modelos hacen un buen trabajo en modelar el comportamiento de compra de los clientes.

Figura 5.

Frecuencia de transacciones repetidas para el modelo BG/NBD

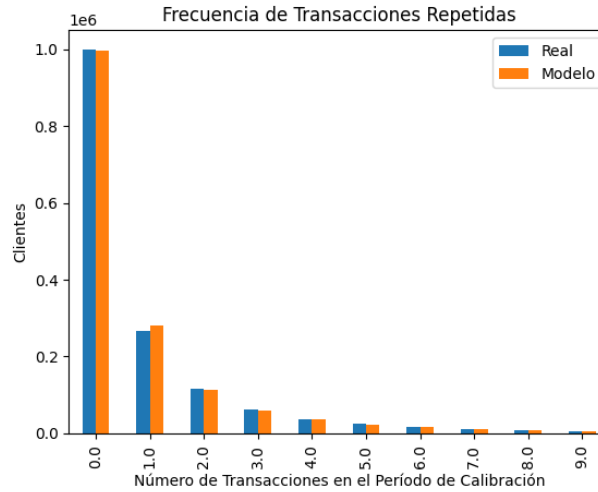
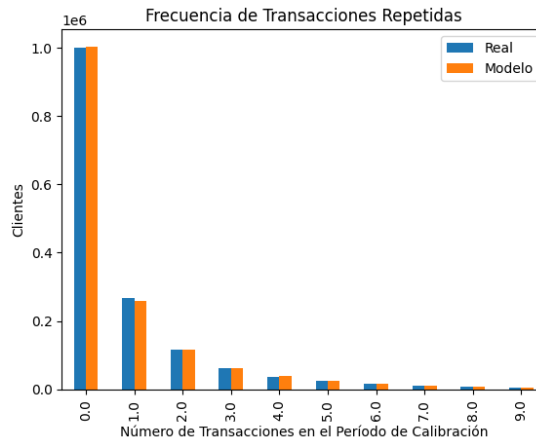


Figura 6.

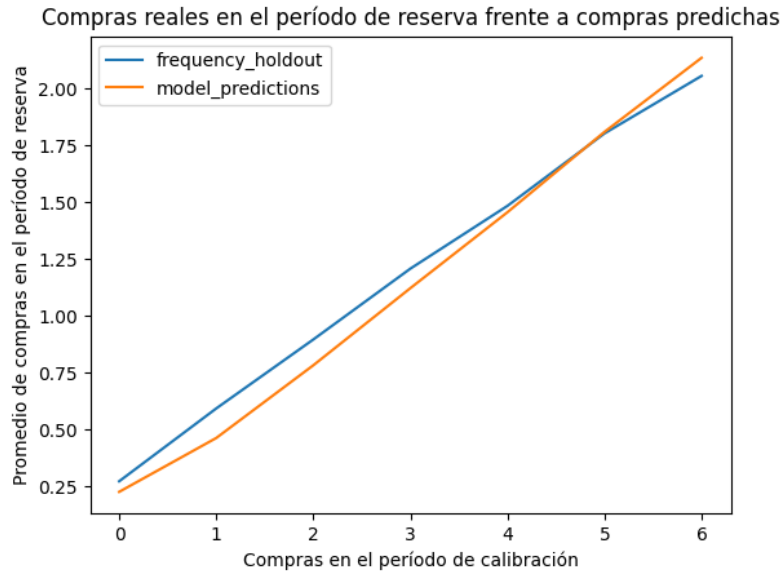
Frecuencia de transacciones repetidas para el modelo MBG/NBD



Para continuar con la validación de nuestros modelos, también es recomendable explorar la relación entre las frecuencias durante el periodo de calibración y las frecuencias reales y predichas durante el periodo de reserva. En la **Figura 7** y **Figura 8** se agrupan los clientes en el periodo de calibración por la cantidad de compras repetidas y luego se promedia sus compras repetidas en el periodo de reserva (Li, 2018). En ambas figuras, se observa que los modelos son capaces de predecir con bastante precisión.

Figura 7.

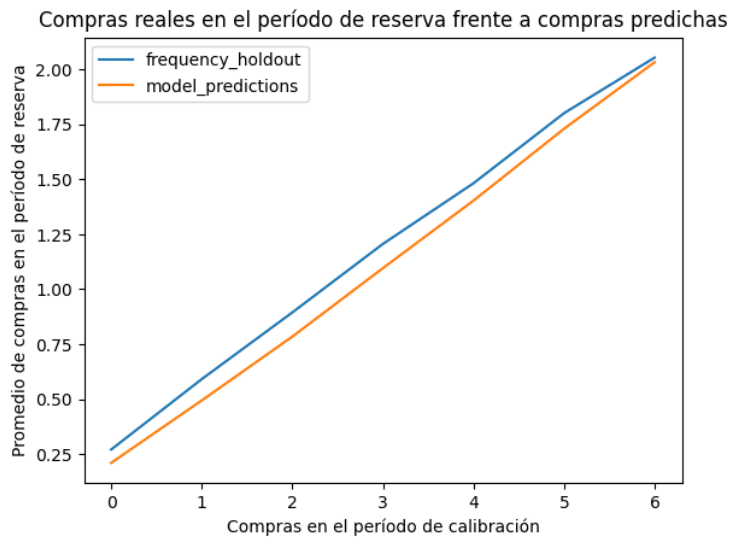
Compras reales en el período de reserva frente a compras predichas (BG/NBD)



Nota. Frequency_holdout se refiere a la frecuencia en el periodo de reserva y model_predictions a las predicciones del modelo. No se cambiaron estos nombres por restricciones de la librería.

Figura 8.

Compras reales en el período de reserva frente a compras predichas (MBG/NBD)

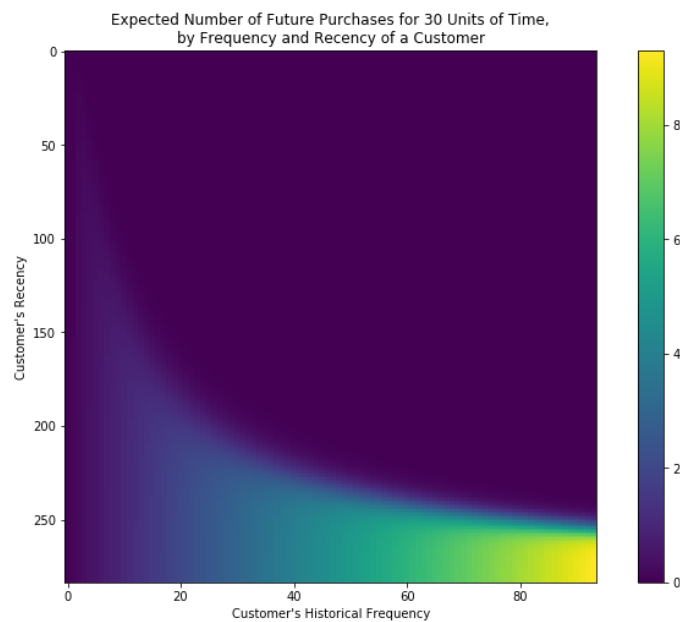


Nota. Frequency_holdout se refiere a la frecuencia en el periodo de reserva y model_predictions a las predicciones del modelo. No se cambiaron estos nombres por restricciones de la librería.

Con un modelo ajustado ya se puede iniciar a hacer algunos análisis en el comportamiento de los clientes. Por ejemplo, en la **Figura 9**, el degradado del color representa la cantidad de compras que un cliente puede realizar en un mes respecto a su recencia y frecuencia. El eje horizontal representa la “frecuencia histórica del cliente”; el eje vertical denota la recencia del cliente. Los clientes más valiosos son los que tienen un alta recencia y frecuencia¹.

Figura 9.

Matriz para analizar las futuras compras respecto a la frecuencia y recencia.



Nota. Matriz futuras compras. [Infografía], Databricks, 2020,
https://www.databricks.com/notebooks/CLV_Part_1_Customer_Lifetimes.html#

Para aplicar el modelo Gamma-Gamma, es necesario verificar que la correlación entre la frecuencia y el valor monetario sea baja, como se muestra en la **Figura 10**. En este caso, se observa poca correlación entre las variables.

¹ Se opta por emplear una imagen externa al proyecto con el fin de preservar la confidencialidad de la empresa.

Figura 10.

Correlación entre las variables frecuencia y valor monetario

	frequency_cal	monetary_value_cal
frequency_cal	1.00	-0.03
monetary_value_cal	-0.03	1.00

Dado que se tiene un modelo para predecir las futuras transacciones de los clientes (MBG/NBD) y otro para calcular el promedio de gasto por transacción, ya es posible estimar el CLV. Para ello se hace uso de la función “customer_lifetime_value”, implementada en la librería Lifetimes, tal como se muestra en la **Figura 11**. Luego de agregar el valor del CLV a nuestra información, tenemos los datos completos para realizar análisis de comportamiento (**Figura 12**), donde tenemos el identificador del cliente con sus valores RFM, las transacciones predichas, el valor promedio por transacción estimado y el Customer Lifetime Value.

Figura 11.

Función para estimar el CLV en un año

```
df_rfmt_copy['CLV'] = (  
    gamma_gamma.customer_lifetime_value(  
        model_MBG,  
        df_rfmt['frequency'],  
        df_rfmt['recency'],  
        df_rfmt['T'],  
        df_rfmt['monetary_value'],  
        time=12,  
        discount_rate=0.0  
    )  
)
```


Figura 12.

Resultado final de la estimación del CLV

	frequency	recency	T	monetary_value	predicted_purchases	monetary_predicted	CLV
SAP							
0000100097	1.00	385.00	2174.00	71798.00	0.06	72003.96	4181.05
0000100104	10.00	1590.00	2272.00	185962.30	0.44	185978.50	81877.01

5 Análisis

En los resultados obtenidos, se evidencia un rendimiento similar entre los modelos empleados, siendo un poco más preciso el modelo MBG/NBD. Esto se puede observar tanto en las diversas gráficas utilizadas para validar los modelos como en el error cuadrático medio resultante de la evaluación del coeficiente de penalización. Es por esto por lo que el modelo MBG/NBD es el seleccionado para realizar los análisis de comportamiento de cliente y calcular el Customer Lifetime Value junto al modelo gamma-gamma elegido.

Por otro lado, debido a la poca información de la empresa con respecto al “discount rate”, un valor crítico que considera la inflación y ajustes de valor futuros, se decidió fijarlo en 0.01, que es el valor por defecto utilizado en la literatura. Cualquier otra cifra que se establezca va a cambiar la estimación, sobre todo a largo plazo (más de 3 años). Con el valor por defecto se detectó una buena estimación para 3 años. Sin embargo, si el *discount rate* se establece en 0.0, no se tendrá en cuenta los cambios de la estimación en base al tiempo, lo que disminuirá su precisión en periodos prolongados.

6 Conclusiones

Se ha logrado implementar con éxito un modelo capaz de estimar el Customer Lifetime Value de los clientes de la empresa. Este modelo facilita el análisis del comportamiento de compra de los clientes, incluyendo la probabilidad de que un cliente esté activo, la cantidad de transacciones que realizará y el monto de dinero que gastará en un periodo determinado.

Los modelos BG/NBD y MBG/NBD demostraron que pueden capturar de manera precisa el comportamiento de compra de los clientes, sin requerir una gran capacidad de cómputo. Por tanto, se perfilan como una herramienta fundamental para cualquier empresa que desee analizar a sus clientes.

Con este trabajo se muestra la importancia de la analítica de datos para las compañías, ya que es una herramienta fundamental en la toma de decisiones. Las empresas cada vez tienen más datos, por lo que es importante conocer diversas prácticas analíticas que faciliten la comprensión de estos datos y su aplicación en la formulación de estrategias que impulsen el crecimiento y la posición competitiva de la compañía.

Referencias

- Batistlam, E. P., Denizel, M., & Filiztekin, A. (2007). Empirical validation and comparison of models for customer base analysis. *International Journal of Research in Marketing*, 24(3), 201-209.
- CLV Part 1: Customer Lifetimes - Databricks. (s. f.). Databricks.com. Recuperado 06 de abril de 2024, de https://www.databricks.com/notebooks/CLV_Part_1_Customer_Lifetimes.html
- Davidson-Pilon, C., Kalderstam, J., Zivich, P., Kuhn, B., Fiore-Gartland, A., Moneda, L., ... & Rendeiro, A. F. (2019). *CamDavidsonPilon/lifelines: v0. 22.4*. Zenodo. <http://doi.org/10.5281/zenodo.3386382>
- Fader, P. S., & Hardie, B. G. (2009). Probability models for customer-base analysis. *Journal of interactive marketing*, 23(1), 61-69.
- Fader, P. S., Hardie, B. G., & Lee, K. L. (2005). "Counting your customers" the easy way: An alternative to the Pareto/NBD model. *Marketing science*, 24(2), 275-284.
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., ... & Sriram, S. (2006). Modeling customer lifetime value. *Journal of service research*, 9(2), 139-155.
- Hu, Y. H., Huang, T. C. K., & Kao, Y. H. (2013). Knowledge discovery of weighted RFM sequential patterns from customer sequence databases. *Journal of systems and software*, 86(3), 779-788.
- Kim, T., Kim, D., & Ahn, Y. (2022). Instant customer base analysis in the financial services sector. *Expert Systems with Applications*, 202, 117326.
- Li, S. (2018). What's a customer worth? Modelling customers lifetime value for non-contractual business with python. *Towards Data Science*. Recuperado el 05 de abril, 2024, de <https://towardsdatascience.com/whats-a-customer-worth-8daf183f8a4f>
- Martins, N. H. (2021). Desarrollo del modelo de Customer Lifetime Value (CLTV) para industria bancaria con técnicas de Machine Learning.
- Montero Montero, M. G., Suárez Villamar, D. J., & Paredes Aguirre, M. (2019). Estimando el customer lifetime value: una métrica alternativa para la segmentación de clientes aplicado al sector retail.
- Mzoughia, M. B., & Limam, M. (2014). An improved BG/NBD approach for modeling purchasing behavior using COM-Poisson distribution. *International Journal of Modeling and Optimization*, 4(2), 141.
- Nikkhahan, B., Badrabadi, A. H., & Tarokh, M. J. (2011). Customer lifetime value model in an online toy store. *Journal of Industrial Engineering International*, 19-31.
- Schmittlein, D. C., Morrison, D. G., & Colombo, R. (1987). Counting your customers: Who-are they and what will they do next?. *Management science*, 33(1), 1-24.

Sobreiro, P., Martinho, D., Pratas, A., Garcia-Alonso, J., & Berrocal, J. (2019). Predicting High-Value Customers in a Portuguese Wine Company. *Journal of Reviews on Global Economics*, 8(2019), 1732-1740.

Zhang, J. Q., Dixit, A., & Friedmann, R. (2010). Customer loyalty and lifetime value: An empirical investigation of consumer packaged goods. *Journal of marketing theory and practice*, 18(2), 127-140.