

Universidad Torcuato Di Tella
Escuela de Negocios
Master in Management + Analytics

Predicción del Net Promoter Score ante la falta de encuestas



Fecha: Mayo 2021

Autor: Franco Santoliquido

Tutor: Ramiro H. Gálvez

Resumen

Para poder retener a los clientes de la compañía es clave que las interacciones con la misma generen buenas experiencias. Una de las principales métricas utilizadas para medir la experiencia de los clientes es el Net Promoter Score (NPS). Este indicador se construye a partir de las encuestas que contestan los clientes. En una plataforma de comercio electrónico, el NPS de los compradores nos permite conocer la experiencia de compra en la plataforma y la *performance* de los vendedores a la hora de transaccionar. Se puede saber si un vendedor está ofreciendo una experiencia satisfactoria a sus clientes y cumpliendo con los estándares pretendidos por la plataforma mirando el NPS que surge de las encuestas que contestan sus compradores. Sin embargo, para garantizar la representatividad y robustez del indicador, es necesario contar con una base mínima de respuestas que no siempre puede obtenerse. En la presente tesis se propone una alternativa de solución ante la falta de encuestas: un algoritmo de *Ridge* que utiliza diferentes variables transaccionales y asociadas a los usuarios para poder predecir el NPS. El algoritmo propuesto demostró tener mejor *performance* que la solución vigente en la empresa. Adicionalmente, permite contar diariamente con un valor de NPS para que los empleados de la empresa encargados de asesorar a los vendedores de la plataforma puedan contar con un KPI que les permita detectar desviaciones en la experiencia ofrecida.

Abstract

It is important for a company to maintain high levels of customer experience in order to retain their customers. One of the main metrics used to measure customer experience is the Net Promoter Score (NPS). This indicator is constructed from the surveys that clients answer. In an e-commerce platform, the NPS of the buyers allows us to know how the sellers are performing and the overall shopping experience. It is possible to know if a seller is offering a satisfactory experience to their customers and complying with the standards set by the platform by looking at the NPS that comes from the surveys that their buyers answer. However, to guarantee the representativeness and robustness of the indicator, it is necessary to have a minimum response base that cannot always be obtained. In this thesis, an alternative solution is proposed in the absence of surveys: a Ridge algorithm that uses different transactional variables to predict the NPS. The proposed algorithm proved to have better performance than the current solution in the company. Additionally, it allows the company to detect whether the sellers are providing the right level of experience by providing it with a daily KPI to follow up.

Índice

1. El problema	5
1.1. Propuesta de trabajo	6
2. Antecedentes	8
2.1. Importancia de la medición del NPS	8
2.2. Construcción del NPS como KPI	10
2. Materiales y métodos	17
2.1. Materiales: datos, encuestas y raw data	17
2.1.1. Los datos de las encuestas	18
2.1.2. Los datos del vendedor	19
2.1.3. Los datos del comprador	21
2.1.4. Los datos de la transacción	22
2.2. Procesamiento de la información	24
2.3. Métodos	27
2.3.1. Métodos de validación y métricas de performance	28
2.3.2. Modelos utilizados	29
3. Exploración de los datos	32
3.1. Los datos del envío	32
3.2. Cancelaciones, categorías y NPS	36
4. Resultados	41
4.1. El modelo base	41
4.2. Combinaciones de variables	42
4.3. Performance de los modelos	44
4.4. Evaluación del modelo	48
5. Discusión	50

5.1. Resumen de resultados	50
5.2. Aplicaciones prácticas	51
5.3. Limitaciones	51
5.4. Conclusión	52
6. Bibliografía	55

1. El problema

Durante las últimas tres décadas, con el desarrollo y expansión de internet, el comercio electrónico se transformó en una de las principales vías de adquisición de bienes y servicios. Al crecimiento orgánico del sector se le suma el contexto de la pandemia de COVID-19 que condujo a muchas personas a realizar sus compras *online*. Tal es el caso, que empresas como Amazon crecieron 38% durante 2020 en ventas o Mercado Libre un 110% en facturación transaccionada durante el mismo período según sus reportes oficiales. Un factor fundamental por el cual las personas eligen y se mantienen operando en estas empresas es la buena experiencia obtenida al transaccionar. En la actualidad, las personas tienen acceso a un mayor nivel de información sobre las empresas, opiniones de otras personas y con ello pueden realizar mejores comparaciones para elegir donde transaccionar. Por lo tanto, las empresas buscan el *feedback* de sus clientes para poder tomar decisiones que mejoren el nivel de satisfacción de los mismos. En el presente trabajo, nos enfocaremos en analizar uno de los indicadores de experiencia y satisfacción de los usuarios de una empresa de comercio electrónico de América Latina: Net Promoter Score (NPS).

Todos los datos utilizados para realizar el trabajo fueron provistos por una empresa de comercio electrónico con operaciones en América. En particular, las operaciones estudiadas, usuarios y conclusiones que se exponen más adelante se limitan a Brasil. El modelo de negocios de la empresa estudiada funciona como una plataforma *multi-sided* (Hagiu, A., 2009). La organización permite la interacción entre dos grupos de clientes diferentes: compradores y vendedores. La participación de la empresa en la operación genera valor a partir de la reducción de los costos de búsqueda de los productos y servicios proponiendo una interfaz web que ayuda a las partes a encontrarse. Además, permite tener costos de transacción menores proveyendo una solución logística y de seguridad en el pago de la compra. Por cada una de las operaciones exitosas que se realicen en la plataforma, el comprador no debe pagar más que el precio del producto. Sin embargo, la empresa cobra una comisión al vendedor que consiste en un porcentaje del total facturado. Este es el principal ingreso de la compañía.

El NPS surge de encuestar a los clientes de la compañía preguntando si recomendarían o no la empresa, basado en su experiencia interactuando con la misma (Reichheld, 2003). Una vez obtenida la respuesta, se sintetiza en un valor numérico con el que se puede obtener un seguimiento en el tiempo para detectar mejoras o no en la satisfacción de los usuarios. Poder evaluar desviaciones de la métrica es importante para poder aplicar acciones correctivas o

impulsar aquellas positivas que generen mejores experiencias. En el contexto de la plataforma de comercio electrónico que se estudiará en este trabajo, el NPS se utiliza como métrica de la *performance* de la experiencia que están brindando los vendedores sobre los compradores. Es decir, se consulta a los compradores luego de haber adquirido un bien de un vendedor de la plataforma cómo ha sido su experiencia de compra. Con todas las encuestas respondidas para las transacciones que realice ese vendedor, se calcula su NPS. Para ello, se computa la diferencia entre la cantidad de valoraciones positivas y negativas y, a ese resultado, se lo divide por la cantidad de encuestas analizadas. El NPS que surge de estas encuestas permite evaluar si el vendedor de esa transacción brinda los niveles de satisfacción esperados por la plataforma.

Un problema que se presenta al momento de calcular el NPS es que el número de encuestas necesarias para obtener un valor representativo de NPS para cada vendedor no siempre es factible de obtener. Los cuestionarios de satisfacción se envían por correo electrónico y la tasa de apertura y respuesta de las encuestas se mantienen en tendencia negativa (Campaign Monitor, 2020). Por lo tanto, resulta necesario encontrar una manera para suplir el faltante de encuestas y obtener un valor de NPS para cada vendedor. El objetivo del presente trabajo consiste en intentar predecir el *NPS de los compradores* todos aquellos vendedores que no cuentan con un nivel suficiente de encuestas para obtener un valor robusto de NPS.

Nuestro aporte será poder obtener el número de NPS de un vendedor utilizando la información disponible en cualquier momento determinado a través de la implementación de un modelo de Machine Learning. Para poder realizar el ajuste del modelo, primero es necesario contar con datos suficientes. En tal sentido, se generará un *dataset* con la información de NPS de cada vendedor y todas las variables transaccionales e históricas que nos permitan predecir ese NPS. La creación de la base implica que utilizaremos encuestas existentes de la empresa y enviaremos nuevas encuestas para alcanzar un valor representativo de NPS para cada uno de los vendedores a analizar. Una vez generada la base, crearemos diferentes variables que permitan predecir el NPS y ajustaremos diversos modelos en busca de aquel con mejores resultados.

1.1 Propuesta de trabajo

La resolución del problema planteado se realiza a través de un flujo que consiste en tres grandes secciones: la creación de la base de encuestas, la generación de las *features* utilizadas

para predecir y el ajuste de diversos modelos en búsqueda de aquel de mejor performance. Los datos recolectados, modelos ajustados y conclusiones están asociadas a una empresa multinacional de comercio electrónico cuyo nombre, por cuestiones de confidencialidad, mantendremos anónimo.

El primer paso, fue la creación de la base en la que se utilizaron las encuestas de los 2085 vendedores más grandes de la plataforma que se encontraban activos durante Octubre 2020. El criterio utilizado para definir que vendedor se encontraba en este top es la cantidad de transacciones vendidas en el último año previo al análisis. Es decir, no se consideran aquellos vendedores que no son usuarios frecuentes o que venden ocasionalmente algún producto, sino a aquellos que generan la mayor cantidad de operaciones e intervienen en la experiencia de la mayor cantidad de compradores.

Una vez obtenidas las encuestas, lo siguiente fue el cálculo del NPS por vendedor y la creación de los *features* con los que se buscó predecir el mismo. Esto supuso la generación de un ETL¹ aplicando el lenguaje SQL a las consultas de múltiples tablas alojadas en un esquema de bases de datos. Completada esta etapa, se exportaron dos archivos “.csv” con un peso de aproximadamente 150 MB cada uno que fueron unidos en Python. Es con el resultado de este *merge* que se obtuvieron las *features* deseadas para el modelo.

Por último, se ajustaron diferentes modelos con distintas combinaciones de *features* para obtener aquel con mejor performance. Luego este último fue comparado contra un modelo básico que es aceptado y utilizado actualmente para poder pronosticar NPS.

El documento consta de 4 capítulos divididos en secciones. En el Capítulo 2, realizamos una revisión de trabajos anteriores y detallamos la importancia de la medición de la experiencia para la compañía de estudio. En el Capítulo 3, describimos los *Materiales y Métodos* utilizados, es decir, el origen de la información y la metodología a seguir. En el Capítulo 4, profundizamos en la exploración de los datos crudos y las variables a usar. En Capítulo 5, explicamos los resultados obtenidos y comparamos los modelos utilizados. Por último, detallamos las conclusiones y aplicaciones del trabajo a la vez que exploramos sus limitaciones y oportunidades para el futuro.

¹ Proceso *Extract Transform and Load* en el que se extraen datos de una base, se limpian y se cargan en otra base para poder ser analizados.

2. Antecedentes

En esta sección se expondrá una revisión de la relevancia que tiene la medición de la experiencia para una compañía y las métricas utilizadas habitualmente haciendo hincapié en el NPS. Luego, detallaremos la manera de calcular este Key Performance Indicator (KPI) así como la metodología asociada a la segmentación y envío de encuestas. Por último, profundizaremos en los procedimientos y problemas que presenta la empresa de estudio para la obtención de las encuestas.

2.1 Importancia de la medición del NPS

En un entorno competitivo, la experiencia de los usuarios al interactuar con la compañía se vuelve un factor fundamental en el crecimiento y futuro de la misma. El hecho de brindar a los clientes una experiencia que supere sus expectativas previas conduce a una fuerte relación entre el cliente y la empresa (Kuo, et. al, 2009). En este estudio nos concentraremos en un KPI que mide la experiencia de los usuarios en el sector de comercio electrónico. El foco estará puesto en el Net Promoter Score (NPS), KPI que surge de un cuestionario de satisfacción de los clientes. Durante el transcurso del trabajo se entiende experiencia del cliente bajo la definición provista por Gentile, Spiller y Noci (2007), donde consideran que es la interacción entre el cliente con el producto, la compañía o una parte de su organización y que genera una reacción estrictamente personal en los niveles racionales, emocionales, sensoriales y físicos.

La posibilidad de acceder a mayor información ha permitido a los clientes poder realizar comparaciones entre productos y empresas a la vez que se ven influenciados por las opiniones del resto en las redes sociales (Almquist et. al, 2018). En consecuencia, mantener altos niveles de satisfacción en los clientes se ha tornado un elemento central en las estrategias de las empresas. Es en este sentido, que las compañías han cambiado sus estrategias de ser *product centric* a *customer centric*. La primera se focaliza en identificar y desarrollar aquellos productos que serán rentables a la compañía y les permitirá ganar *market share* (Shah et al. 2006). En la visión *customer centric* las decisiones de la empresa están orientadas a servir a los clientes, conocer sus inquietudes para generar vínculos a largo plazo (Fader, 2020). El objetivo final es incrementar y mantener una elevada lealtad de los clientes para evitar que los mismos migren a la competencia. En la literatura de Marketing, se ha argumentado que el costo de atraer nuevos clientes es significativamente mayor al de retener a una base ya creada, llegando el primero a

ser cinco veces el segundo (Pfeifer, 2005). Dentro de los costos de atracción de clientes se incluyen aquellos llamados “*land grab*” (Pei-Yu et. al, 2002), que envuelven a las campañas de *Search Engine Optimization*, promociones y subsidios a las primeras compras. Ejemplos de estos gastos son los de Rappi y Magazine Luiza, quienes, en Brasil, durante 2020 ofrecieron cupones de descuento en la primera compra o *cashback* en sus billeteras virtuales. Uno de los métodos utilizados para evitar que los usuarios migren hacia la competencia es el *churn management* (Hung, et. al., 2006). Las empresas crean gerencias específicas para entender cuánto valor les aporta cada uno de sus clientes a la vez que crean modelos que les permitan identificar de manera anticipadas quienes de ellos tienen una tendencia mayor a abandonar la empresa.

En el caso de una empresa de e-commerce la experiencia podemos dividirla en tres etapas: la búsqueda, la compra y la post-compra. En la etapa de búsqueda, el usuario tiene su primer encuentro con la compañía, la cual será considerada exitosa en tanto pueda encontrar el producto que está buscando. En un estudio, la compañía Nielsen (2000) testeó 20 sitios web con usuarios intentando determinar el momento en el que se cae la compra. Se destaca la importancia de los diseñadores y desarrolladores web en los sitios de e-commerce bajo el argumento de que el 27% de los usuarios testeados se caen por dificultades a la hora de buscar. Una vez que el usuario encuentra el producto deseado, ingresa en la etapa de compra. Aquí entran en juego las variables como el precio, los métodos de pago disponibles, los tiempos y costo de envío sumado a la confianza que se tiene en los vendedores. Uno de los métodos que utilizan las empresas de comercio electrónico para generar confianza en los vendedores de su plataforma son las reseñas de otros compradores. Estas *reviews* le permiten a los usuarios tener un panorama de cómo será su experiencia comprando a ese vendedor y qué puede esperar del producto buscado. Un punto importante a mencionar es que al ser independientes de la plataforma y no tener intereses compartidos, los usuarios asignan mayor credibilidad a las mismas (Cheong H., Morrison M., 2008). Mayores y mejores *reviews* se traducirán en mayores ventas de los productos de estos vendedores (Davis A., Khazanchi, D., 2008). La última etapa es la post-compra, en la que el usuario recibe su producto. El usuario interactúa con la compañía dando una reseña de su compra, contestando una encuesta de satisfacción o contactando con atención al cliente ante cualquier problema que tenga con su compra (esto es, demoras, productos diferentes o defectuosos). A lo largo de este estudio, nos concentramos en la segunda etapa del proceso de compra de los usuarios de una empresa de comercio electrónico.

2.2 Construcción del NPS como KPI

Con el objetivo de medir la experiencia de los clientes con la empresa, en la industria de retail y tecnológica se utilizan una serie de modelos y KPIs que guían las decisiones de los managers. Según el estudio Bain, durante el último tiempo se ha acelerado el uso de los modelos de Customer Lifetime Value (CLV), Propensity model, Sentiment Analysis y Net Promoter Score (NPS) en la gestión de las decisiones gerenciales de la empresa (Du Toit et. al, 2020).

Esta última métrica, ha sido adoptada por muchas empresas debido a que se origina en una única y simple pregunta: “*basado en tu última experiencia con la compañía ¿qué tan probable es que la recomiendes a un amigo o colega?*” (Reichheld, 2003). El cliente elige en una escala del 0 al 10 siendo 0 “nada probable” y 10 “muy probable”. Según la metodología propuesta por Reichheld (2003), aquellos que eligen en la escala 9-10 son considerados “Promotores”, quienes elijan 7-8 son “Pasivos” y, por último, aquellos que escojan entre 0-6 son “Detractores”.



Figura 1: Distribución de Promotores, Pasivos y Detractores. Fuente: elaboración propia en base a Reichheld 2003.

Los Promotores, tienen una percepción positiva de la empresa y recomendarán a la empresa a través del “boca a boca” (*word of mouth*). Es a través de este grupo de personas satisfechas que la empresa no sólo mantiene una base de clientes que realizará futuras compras, sino que ayudará a expandir el *share* de clientes a través de la recomendación. Por su parte, los “detractores”, quienes están disconformes con sus interacciones con la empresa, no recomendarán sus servicios y al hablar con potenciales nuevos clientes que pertenezcan a su entorno, los convencerán de no consumir los productos o servicios de la compañía. En este sentido, sostener un conjunto grande de “Detractores” conducirá a reducir la base de clientes y afectar el crecimiento futuro. Por último, los usuarios “Pasivos” son indiferentes respecto a la recomendación y no son considerados como leales a la empresa. Esto implica que son pasibles

de ser persuadidos ante una mejor oferta y se moverán hacia la competencia (Reichheld, 2003, 2006). Una de las finalidades del seguimiento del NPS es entender y resolver cuáles son los problemas que tienen los Detractores para transformarlos en Pasivos o Promotores, accionar sobre las inquietudes de los Pasivos y reforzar aquellos puntos que los Promotores valoran sobre sus interacciones con la compañía.

Para calcular el KPI final y sintetizar en un único valor la experiencia del usuario se sustrae de la cantidad de Promotores a la cantidad de Detractores y se los divide por el total de personas encuestadas. De esta manera, se obtiene un valor continuo acotado entre -1 (en el caso de que el 100% de los encuestados sean Detractores) y 1 (en el caso de que todos los encuestados sean Promotores). Para obtener un valor representativo de NPS, manteniendo un margen de error de 10%, las encuestadoras sugieren que para una población objetivo por encima de 10.000 clientes, el número de personas a encuestar debe ser de al menos 97. En este estudio intentaremos generar un modelo que nos permita estimar ese valor de NPS cuando las encuestas no son suficientes.

$$NPS = \frac{(\#Promotores - \#Detractores)}{Total\ encuestados} [-1, 1]$$

Según Reichheld (2003), la pregunta de NPS es el único recurso que las empresas deben utilizar para medir la lealtad de sus clientes. Sin embargo, menciona que se puede realizar una pregunta de seguimiento para poder entender el motivo detrás de la calificación. En esta pregunta se le solicita al usuario que indique cual es el principal motivo por el cual dió la calificación del 1-10 o que seleccione de un listado de opciones todos aquellos factores que influyeron en su decisión. Como último punto, algunas empresas consultoras² que ofrecen su plataforma para crear encuestas de NPS recomiendan dejar una pregunta abierta donde el cliente pueda expresar y dar mayor contexto sobre el puntaje dado. En las imágenes 1, 2 y 3 se pueden observar ejemplos de encuestas de NPS ofrecidos por las empresas SurveyMonkey.com y Hotjar.com.

² Qualtrics, SurveyMonkey, Hotjar

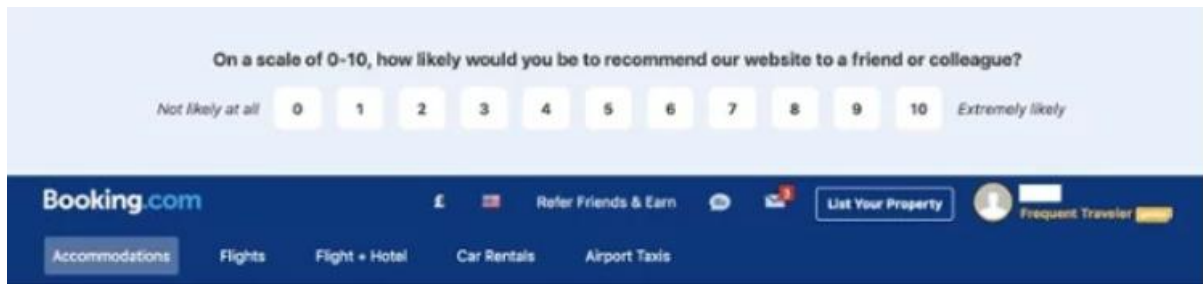


Imagen 1. Ejemplo de pregunta de NPS en *Booking.com* como banner en página principal.

Fuente: www.hotjar.com/net-promoter-score

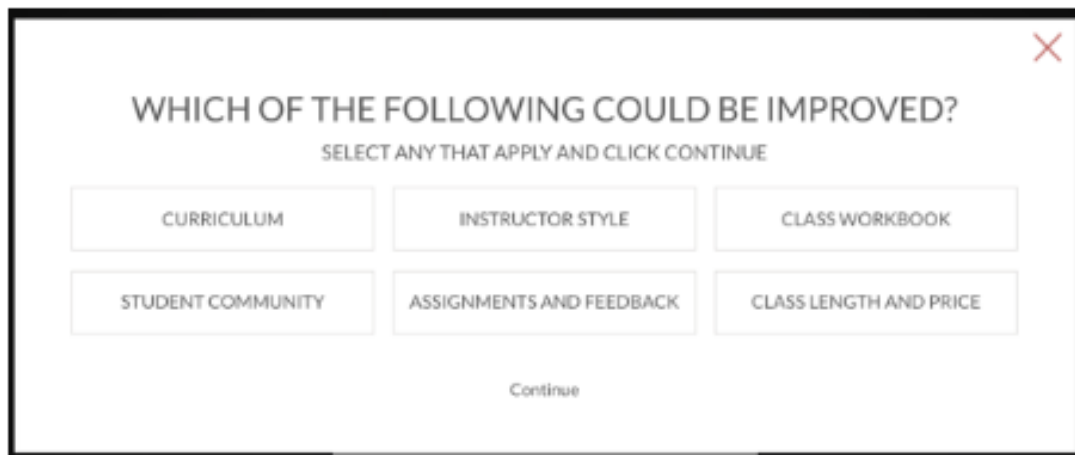


Imagen 2. Ejemplo de pregunta de seguimiento Fuente: www.hotjar.com/net-promoter-score

Feedback for Improvement

2. What changes would this company have to make for you to give it an even higher rating?

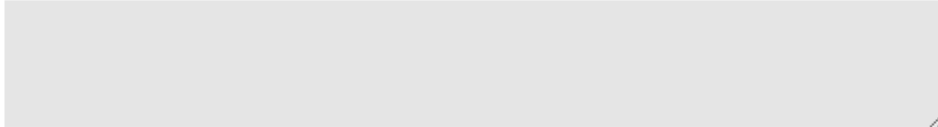


Imagen 3. Ejemplo de pregunta abierta de seguimiento. Fuente: www.surveymonkey.com

El estudio original de Reichheld (2003) obtuvo datos de 14 compañías que pertenecían a diferentes industrias. Luego de realizar encuestas para esas empresas, calculó el valor de su NPS y afirmó que este indicador tenía una mejor performance a la hora de predecir el comportamiento de compra de los clientes. En estudios posteriores, menciona la existencia de una fuerte correlación entre el NPS y las tasas de crecimiento de las compañías (Reichheld, 2006). Es en este sentido que Reichheld (2003, 2006) afirma que la pregunta de NPS es lo *único* que las empresas deben preocuparse por preguntar a sus clientes siendo el indicador central que deben mantener alto.

Sin embargo, en la literatura de gestión organizacional, el NPS ha recibido ciertas críticas como medida de lealtad de los clientes. Los argumentos en contra del uso de la métrica están asociados principalmente a considerarla como el único valor que las empresas deben seguir (Zaki et. al, 2016). Una única métrica de lealtad no debe guiar las decisiones de los gerentes porque se considera que los comportamientos de los usuarios son multidimensionales y no pueden ser capturados mediante un único KPI. Adicionalmente, se afirma que la utilización aislada del NPS no brinda una respuesta clara a los factores causales de las experiencias positivas o negativas de los clientes y que es necesario conducir investigaciones adicionales para entender estos factores (Grisaffe, 2007). Por ejemplo, la práctica recomendada es utilizar reseñas escritas por los usuarios para entender por qué dejan de comprar o por qué siguen comprando. Es en este sentido, que la recomendación a la hora de estudiar la experiencia de los usuarios de manera integral y poder tomar mejores decisiones es considerar un conjunto

de modelos y herramientas, evitando limitar los estudios sobre lealtad y acciones de marketing de las empresas al movimiento de una sola métrica (Zaki et. al, 2016).

Siendo parte de un conjunto de métricas, el NPS es de gran utilidad debido a su facilidad de comprensión y funcionamiento de “termómetro” de la performance de la compañía. En la empresa de comercio electrónico estudiada, la información sobre NPS se obtiene de encuestas enviadas por correo electrónico. Luego de realizada una compra, el usuario puede ser elegido para responder la encuesta mediante un muestreo aleatorio estratificado. Para generarlo se definen los segmentos de usuarios a encuestar teniendo en cuenta el método de envío utilizado en la compra, el tamaño del vendedor y el nivel de antigüedad o lealtad del comprador. Para cada segmento, se establece un número de encuestas a enviar en función de la cantidad de compras que se realizaron en cada combinación método de envío - tamaño del vendedor - tipo de comprador. Una vez definido el tamaño de encuestas a enviar se seleccionan aleatoriamente un conjunto de usuarios a los que se les enviará la encuesta.

Es factible que un usuario que realiza muchas compras en un período corto de tiempo sea seleccionado aleatoriamente por el modelo varias veces para el envío de la encuesta. Para evitar molestar al usuario con tantas encuestas seguidas o que los correos electrónicos lleguen a la casilla de “spam”, se establece un tiempo de al menos 15 días entre cada encuesta respondida por el usuario. Es decir, si un cliente recibe y responde una encuesta en un mes, se esperan al menos 15 días hasta enviarle otra encuesta.



Imagen 3. Flujo de envío de encuestas de NPS. Fuente: Elaboración propia

El problema detectado en el procedimiento descrito para la empresa estudiada es la tasa de respuesta de las encuestas enviadas vía correo electrónico³. Abrir el correo y navegar para

³ Una alternativa es modificar el método de envío y hacerlo vía notificaciones push o modal. Sin embargo, todas las áreas de la compañía (i.e., Marketing, Legales, Seguridad Informática) utilizan este método para sus

llegar al *link* que redirige hacia la encuesta, supone un esfuerzo adicional para los usuarios. Al mismo tiempo, el mail es una de las principales vías de fraude por internet (*phishing*): los correos son similares a las instituciones reales y solicitan información sensible de los clientes como cuentas bancarias o contraseñas (Aburrous et al., 2010). Ambos factores conducen a una menor tasa de apertura de e-mails en las industrias de comercio electrónico, bancaria y retail. Según un reporte de la empresa de software Campaign Monitor (2020⁴) -especializada en la creación y administración de e-mails de Marketing- la tasa de apertura de los correos electrónicos en el sector de Retail fue de un 7.76% pasando de 14.98% en 2019 a 13.9% durante 2020, de 15.48% a 13% en Alimentos y Bebidas y de 17% a 12.6% en el sector Automotor. En consecuencia, contar con un modelo que permita predecir NPS utilizando otras variables por fuera de las encuestas es importante en un contexto de reducción de tasas de respuesta de las mismas.

Para la empresa de comercio electrónico estudiada, la medición del *NPS de los compradores* ayuda a entender la calidad del servicio que ofrecen los vendedores y la plataforma a la hora de transaccionar. Todas las transacciones que se realizan en un sitio de comercio electrónico que actúa como plataforma intermediaria repercuten en la imagen de la empresa. Es importante que los vendedores, principalmente aquellos que mueven grandes volúmenes de órdenes, mantengan altos estándares de calidad porque el responsable final del éxito de la transacción es la empresa intermediaria. Si esta última no se asegura de que los vendedores son confiables y que las operaciones realizadas son seguras, los compradores pueden migrar hacia la competencia. En términos de NPS, los promotores se convertirán en neutros, los neutros en detractores y se perderán los beneficios de la recomendación y la recompra. El seguimiento de esta métrica ayuda a los empleados de la plataforma que actúan como asesores comerciales de los grandes vendedores a diferenciar cuáles de ellos están ofreciendo una buena experiencia a los compradores de aquellos que tienen una performance pobre. Pueden utilizar como alarma un KPI comprensible que sintetiza todo un conjunto de métricas de experiencia. Cuando el NPS se desvía de los valores habituales o permanece por debajo de aquel de otros vendedores comparables, los asesores comerciales pueden comunicarse con las respectivas empresas para plantear la mejor estrategia para mejorar este indicador. El objetivo del trabajo será predecir el *NPS de los compradores* para aquellos

comunicaciones por lo que el espacio en las notificaciones es limitado. Se busca evitar el envío de múltiples notificaciones para no afectar a la usabilidad de la plataforma.

⁴ <https://www.campaignmonitor.com/resources/guides/email-marketing-benchmarks/>

vendedores asesorados para los que la empresa no tiene encuestas suficientes utilizando un conjunto de variables.

3. Materiales y métodos

En este capítulo se explicará en detalle cuáles fueron los datos utilizados, la metodología de transformación y validación de datos adoptada y, finalmente, los modelos utilizados. Para poder realizar el estudio, debemos crear un *dataset* con el NPS de cada vendedor a estudiar. Luego, ajustar diferentes modelos y quedarnos con aquel que tenga mejor performance para predecir el NPS. Por lo tanto, necesitamos una cantidad suficiente de encuestas por vendedor para luego calcular su NPS. A fin de resolver esto, utilizamos encuestas existentes y realizamos un envío manual para completar la cantidad faltante en ciertos vendedores. Por otro lado, para completar el dataset y obtener las *features* del modelo, debemos contar con datos sobre las transacciones, los envíos de paquetes y el historial de ventas y compras.

3.1. Materiales: datos, encuestas y raw data

El análisis fue realizado en Octubre 2020, por lo que el proceso de creación del dataset se basó en la utilización de encuestas existentes enviadas periódicamente entre 2019 y Septiembre 2020 bajo el esquema de segmentación y envío automático: el usuario compra, se segmenta, se muestrea y se envía la encuesta. Además se realizó un reenvío adicional de encuestas que no fueron contestadas para elevar la tasa de respuesta obtenida originalmente en el proceso automático. Esto implicó consultar cuáles de las encuestas ya enviadas no fueron respondidas y generar un envío manual con un recordatorio de respuesta. Esto implicó el envío de 160.000 encuestas más. En tal sentido, buscamos obtener la mayor representatividad posible, garantizando al menos 97 encuestas por vendedor. Con ese objetivo, la cantidad total de encuestas utilizadas para ser agrupadas en los 2085 vendedores fue de 220.098, obteniéndose un promedio de 105 encuestas por vendedor y un mínimo de 97 encuestas en aquel con menor cantidad.

Para cada respuesta obtenida buscamos toda la información que acompaña a la misma y que luego será utilizada para crear los *features* del modelo. Esta información puede ser dividida en cuatro grupos: los datos de las encuestas, los datos del vendedor, los datos del comprador y los datos de la operación.

3.1.1. Los datos de las encuestas:

Los datos de las encuestas se encuentran asociados a un identificador que llamaremos “encuesta_id” a través del cual podemos identificar el puntaje brindado por el usuario (PUNTAJE en la tabla 1), el motivo de promoción o detracción según corresponda dado el puntaje dado por quien responde (M_PROMOCION, M_DETRACCION) y la fecha de respuesta (RESP_FECHA). Adicionalmente, se cuenta con un identificador tanto de la operación en cuestión por la que se encuesta al usuario (OPER_ID) como otros dos para localizar a quien responde la encuesta (COMP) cómo a quién vendió el producto (VEND). A continuación mostramos un ejemplo de la disposición de los datos de la encuesta.

Tabla 1: Muestra⁵ de la información obtenida en las encuestas

ENCUESTA_ID	RESP_FECHA	PUNTAJE	M_PROMOCION	M_DETRACCION	OPER_ID	COMP	VEND
19203	2019-11-25	9	Plazo entrega		2203560	163189	205852
251934	2020-06-01	8			2439197	161847	700551
28962	2020-08-23	6		Demora en la entrega	262982	144414	857019
24340	2020-05-01	6		Insatisfecho con producto	239679	800349	33220883
36460	2020-07-07	10	Plazo entrega		2530518	571329	25126949
171890	2019-09-04	10	Buen precio		2122741	298799	25081156
21080	2020-02-12	1		Demora en la entrega	2278321	140925	18610110
302	2020-10-24	10	Buen precio		4090597	420399	28768697

⁵ Los identificadores de operación, comprador y vendedor fueron truncados dos dígitos para mantener la confidencialidad de los usuarios

12215	2019-02-09	10	Buen vendedor		1923502	143844	37710456
2642	2020-07-10	10	Buen precio		2520217	374622	21675732
2362	2020-04-05	8			2369667	369247	12698419
18994	2019-10-25	5		Precio	2176846	226901	26499625
5932	2020-01-25	8			2279488	613984	20635561
2229	2020-09-17	10	Otros motivos		4020425	420302	18742725

3.1.2. Los datos del vendedor:

Contando con el identificador del usuario con el rol de vendedor en la operación por la que se encuesta al comprador podemos obtener información sobre sus características y comportamientos pasados. Es decir, exploramos qué tipo de productos vende, cuantos items vendió en el período analizado y qué tan exitosas fueron sus operaciones pasadas. En este sentido, detallamos a continuación cada una de las variables que obtuvimos para crear los features asociados al vendedor:

- **Segmento por nivel de ventas (SEGM):** Segmento asociado según su nivel de facturación y cantidad de ventas. Dado que es una variable que puede cambiar en el tiempo, utilizamos el existente al momento del análisis. Tiene cuatro categorías LIDERES UP, LÍDERES MID, MIDTAIL, LONGTAIL
- **Semáforo (SEMA):** categoría brindada según la experiencia brindada, quejas y comportamiento de entrega. Es una variable que puede cambiar en el tiempo, utilizamos el existente al momento del análisis.

- **Categoría de ventas (CATEG):** Tipos de producto que vende el usuario. Es una variable categórica con 6 valores posibles: electrónica, hogar, supermercado, vehículos y accesorios, moda, otros.
- **Cantidad de ventas mensuales (Q_MONTH1):** Cuántas operaciones realizó el vendedor mensualmente en los últimos 1, 3, 6 meses.
- **Facturación mensual (FACT_M_N):** Cuánto dinero generó el vendedor diariamente en los últimos 30, 60 y 180 días.
- **Cantidad de reclamos (REC_N):** es una variable que nos permite conocer si existió algún problema con la transacción realizada. Pueden ser reclamos por venta de productos deficientes o rotos, perdidos en viaje o demorados en el envío. Cada operación puede tener o no un reclamo asociado, por lo tanto lo que hacemos es calcular para un fragmento temporal y relativo al total de operaciones, cuántas tuvieron un reclamo asociado. Se calcula para diferentes frecuencias temporales: cuantos reclamos tuvo el vendedor en los últimos 5, 15, 30 días sobre el total de operaciones realizadas en ese período.
- **Cantidad de operaciones canceladas (CANC_N):** similar a los reclamos, nos permiten conocer si fue o no una operación exitosa. Son ventas que se concretan pero luego se cancelan antes de que el paquete llegue a destino. Puede deberse a que el vendedor se quedó sin stock, el comprador se arrepienta de la compra, el envío esté demorado o por seguridad si alguna de las partes no está conforme con el comportamiento de la otra. Esta variable también se calcula en diferentes frecuencias temporales: cuántas operaciones fueron canceladas en los últimos 5, 15, 30 días sobre el total de operaciones realizadas en ese período.

Tabla 2: Muestra⁶ de la información asociada al vendedor

VEND_ID	SEGM	SEMA	CATEG	Q_MONTH1	FACT_M30	REC_30	CANC_30
205852	LÍDERES MID	green plat	CE	2470	45882.99	243	210
700551	LIDERES UP	green	HOME	23634	1694633	2119	1796

⁶ Los identificadores del vendedor fueron truncados dos dígitos para mantener la confidencialidad de los usuarios

857019	LÍDERES MID	green plat	ACC	3900	48084.39	335	289
33220883	LÍDERES MID	green gold	HOME	2318	171346	132	117
25126949	LIDERES UP	green	ACC	3296	127056	218	202
25081156	LIDERES UP	green plat	HOME	427	21336.83	24	19
18610110	LIDERES UP	green plat	CE	669	16722.70	32	28
28768697	LIDERES UP	green plat	HOME	60606	2798575	6306	2715
37710456	MIDTAIL	green plat	APPAREL	20796	1466146	1696	1458
21675732	MIDTAIL	green plat	APPAREL	3607	28569	490	383
12698419	MIDTAIL	green plat	HOME	2044	66759	170	131
26499625	LIDERES UP	green plat	ACC	3369	100655	384	357
20635561	LIDERES UP	green plat	ACC	4539	135135	290	246
18742725	LIDERES UP	green plat	APPAREL	2156	16722.70	122	119
556231	LIDERES UP	green plat	CE	6420	158392	317	212

3.1.3. Los datos del comprador:

Utilizando el campo de la encuesta que permite reconocer al comprador, obtenemos los siguientes datos para sumarlos como posibles *features*:

- **Segmento de frecuencia de compra (SEGM_COMP):** es el grupo al que pertenece el usuario según la regularidad de sus compras. Pueden ser usuarios nuevos, usuarios regulares, usuarios frecuentes.
- **Cantidad de compras (Q_COMPRAS_N):** cantidad de compras efectuadas en el último año, en los últimos 3 y 6 meses y en el último mes.

Tabla 3: Muestra de la información asociada al comprador

COMP	SEGM_COMP	Q_COMPRAS_3	Q_COMPRAS_6	Q_COMPRAS_12	Q_COMPRAS_1
4589298	FREQUENT_BUYER	7	7	9	7
292684569	STANDARD_BUYER	4	5	5	2
282258162	FREQUENT_BUYER	5	8	12	0
24561951	NEW_BUYER	0	1	0	0

3.1.4. Los datos de la transacción:

Este segmento de fuentes de datos se obtiene a partir del identificador único de la transacción. A diferencia de los datos del comprador y del vendedor, este bloque de datos manifiesta la experiencia puntual de la operación. Esto es, podemos entender si la transacción que luego genera una encuesta de NPS fue exitosa generando una buena experiencia puntual o si existió alguna fricción que pudo haber impactado negativamente en la experiencia de compra. Teniendo en cuenta esta perspectiva, se obtuvieron datos sobre los días de demora en la entrega, el costo del envío, el precio del producto combinado con la cantidad de productos comprados, el método utilizado para enviarse y los dos *flags* de éxito de la operación cancelación y reclamo asociado. A continuación se describe el detalle de las variables obtenidas.

- **Días de entrega (D_ENT):** cantidad de días desde que el usuario compró el producto hasta que fue entregado.

- **Días de demora (D_DEM):** cantidad de días desde la promesa de entrega hasta que fue efectivamente recibido por el usuario. Mantiene un valor positivo si hubo demoras en la entrega, cero si fue entregado el día prometido y negativo si fue entregado previo a la promesa. Si la operación fue cancelada no se calcula.
- **Cantidad de ítems (Q_ITE):** Responde a la cantidad de elementos comprados por el usuario en la operación. Por ejemplo, un usuario puede comprar dos celulares iguales en la misma operación.
- **Costo del envío (ENV_COST):** es el precio pagado por el comprador correspondiente al flete de entrega. Si la operación tiene envío gratuito el costo es cero. Se expresa en moneda local.
- **Estado de la operación (ESTADO):** si la operación fue exitosa, cancelada o el producto devuelto.
- **Método de pago utilizado (MET_PAY):** esta variable permite entender si la operación fue pagada con tarjeta de débito, tarjeta de crédito, dinero digital, efectivo en un centro de pago habilitado.
- **Método de envío (MET_ENV):** permite conocer si el paquete viaja por el correo tradicional enviado por el vendedor, si permanece en un centro de distribución contratado por la plataforma o si es enviado a un punto de retiro y recolectado por el comprador.
- **Gasto total de la operación (GASTO):** corresponde al precio de cada uno de los elementos comprados en la operación por su precio. Más adelante explicaremos cómo asociar esta variable al costo de envío para entender el porcentaje que representa el envío en el total del gasto.
- **Concentración del mercado (CONCE):** esta variable busca capturar la cantidad de opciones que tiene el comprador dentro de la plataforma para el producto que está buscando. Se calcula la concentración del mercado para cada categoría de venta para el día en el que el comprador realiza la operación.

Tabla 4: Muestra de la información asociada a la orden

OPER_ID	ESTADO	D_ENT	D_DEM	Q_ITE	ENV_COST	MET_PAY	MET_ENV	GASTO	CONCE
4107686	exitosa	6.0	delay	1	23.9	ticket	correo	14.33	3476
1983179	exitosa	2.0	On time	1	12.9	ticket	correo	7.73	118
2221426	exitosa	3.0	On time	1	8.9	c_card	centro_2	13.06	670
2062379	exitosa	11.0	Delay	1	0.0	c_card	centro_2	58.82	556
2209069	exitosa	2.0	On time	1	7.9	c_card	correo	20.79	522
1977804	exitosa	1.0	On time	4	0.0	c_card	centro_1	42.40	463
2146831	exitosa	2.0	On time	1	26.9	c_card	centro_2	19.17	225
4135433	exitosa	1.0	on_time	1	8.9	c_card	centro_1	9.53	799
4090356	cancel	3.0	on_time	1	6.9	acc money	centro_2	5.60	186.0
2417501	exitosa	2.0	on_time	1	30.9	c_card	centro_1	47.25	327
2268890	exitosa	1.0	Early	1	17.9	c_card	centro_2	17.24	486
2561477	DevEx	1.0	early	1	0.0	c_card	centro_2	126.87	522
4047271	exitosa	3.0	on_time	2	23.9	ticket	centro_2	23.84	2500
4050714	exitosa	1.0	on_time	1	22.9	c_card	centro_1	10.94	164

3.2. Procesamiento de la información

Con el objetivo de obtener mejores features que permitan predecir el NPS de los compradores, se realizó un procesamiento de la información. El mismo consiste en combinar, modificar o cambiar la escala de las variables crudas detalladas en la sección anterior

obteniendo así nuevas features. En primer lugar, utilizando la *información asociada a la orden* descrita en la sección 3.1.4, se creó la variable “**cost_percent**” que intenta capturar cuánto del gasto total realizado en la compra se debe al costo de envío. La fórmula aplicada es

$$cost\ percent = \frac{(Costo\ de\ envío\ en\ moneda\ local)}{Total\ gasto\ en\ moneda\ local}$$

Uno de los motivos de detracción de los usuarios que se asocian a la encuesta es el costo del envío. Al expresarlo como proporción del gasto de la compra buscamos capturar el efecto en la experiencia de gasto en flete demasiado caro para un producto de precio bajo.

La segunda variable generada con *información asociada a la orden* es una variable dicotómica que indica si la operación fue entregada a tiempo o si fue entregada luego de la promesa. En función de esto, se calculó el porcentaje de envíos entregado en el tiempo prometido que tuvo el usuario en estas órdenes de la siguiente manera:

$$demora\ percent = \frac{(Órdenes\ entregadas\ a\ tiempo)}{Total\ órdenes\ entregadas}$$

Siguiendo una lógica similar, se construyeron dos variables adicionales: “cancel_percent” y “claim_percent”. La primera busca capturar del total de órdenes analizadas, qué porcentaje terminó siendo cancelada por alguna de las partes. Por su parte, la segunda intenta capturar aquellas que tuvieron algún inconveniente que derivó en un reclamo.

Por otro lado, buscamos crear una variable que permita entender la concentración del mercado a la hora de elegir un producto dentro de una categoría. El objetivo de esta variable es poder capturar el poder de decisión del comprador y la cantidad de alternativas que tiene para encontrar su producto. En un mercado muy concentrado con pocas opciones de vendedores con distintos precios ofrecidos para el mismo producto puede jugar en contra de la experiencia de búsqueda y posterior compra. Un método utilizado comúnmente en trabajos de economía y derecho (Whinston, 2006; Viscusi et al., 2005) para calcular la concentración del mercado es

el índice de Herfindahl-Hirschman (HHI). Puede tomar valores entre $(1/N)$ y 10000 si el mercado es monopolístico y una sola empresa vende todos los productos. Su fórmula se expresa de la siguiente manera:

$$\sum_{i=1}^N S_i^2$$

Donde S es la participación de la empresa en el mercado. En nuestro trabajo, consideramos al “mercado” como la subcategoría en la que compra un usuario. Esto es, si busca comprar un celular, su mercado será la subcategoría⁷ de compra “celulares”. Siguiendo esta fórmula, se creó la variable HH_INDEX siguiendo esta serie de pasos:

1. Por día y por cada subcategoría (ejemplo celulares, lavadoras, lámparas, zapatillas deportivas, etc.) se calcula la cantidad de ventas totales y la participación de cada uno de los vendedores.
2. La participación de cada vendedor se eleva al cuadrado y se calcula el HHI por día por subcategoría.
3. Para cada transacción analizada y asociada a una encuesta, buscamos a qué subcategoría pertenece el producto comprado y le asociamos el HHI de esa categoría en el día de la compra.
4. Por último hacemos una agregación a nivel vendedor tomando el promedio y el desvío estándar de los HHI de las órdenes vendidas por ese usuario.

Los últimos dos métodos de transformación de los datos crudos son la creación de variables dummy para determinar la categoría de venta de un vendedor y la modificación de la escala de variables numéricas. Para la primera, se calcularon las unidades de ventas de los usuarios en cada una de las principales categorías y se le asignó al vendedor aquella categoría en la que más ventas tuvo. Esto generó una variable con siete categorías distintas. Mediante la

⁷ Una subcategoría se considera al primer nivel de desagregación de las 7 categorías principales. Siendo una categoría principal “electrónica”, sus subcategorías serían “celulares”, “lavarropas”, “heladeras”, “laptops”, “computadoras”, “televisores”, “hornos y fogones”, “microondas”, “consolas”, “otros”.

técnica One Hot Encoding creamos seis (utilizando cinco en el modelo) nuevas variables para el modelo que toman valores 1 y 0 si el usuario vende en esa categoría:

1. `home_dummy`: toma valor 1 si la principal categoría de venta es artículos para el hogar
2. `acc_dummy`: toma valor 1 si la principal categoría de venta es accesorios para automóviles
3. `app_dummy`: toma valor 1 si la principal categoría de venta es vestimenta
4. `ce_dummy`: toma valor 1 si la principal categoría de venta es electrónica
5. `cpg_dummy`: toma valor 1 si la principal categoría de venta es artículos de supermercado
6. `other_dummy`: toma valor 1 si la principal categoría no corresponde a hogar, accesorios para automóviles, vestimenta, electrónica ni artículos de supermercado.

Por último, dado que varios de los modelos que se buscan testear en este trabajo se pueden ver afectados por la escala de las variables, se decidió hacer una transformación escalando las variables a través del escalador MinMax. Una de las ventajas de esta técnica es que la distribución de los datos no se ve alterada por la aplicación de la misma. Esta transformación permite modificar la escala de las variables pasando al entorno [0, 1], siendo su fórmula:

$$X = \frac{x - \min(x)}{\max(x) - \min(x)}$$

3.3. Métodos

El objetivo fue predecir una variable continua (NPS) por lo tanto el tipo de problema al que nos enfrentamos fue de regresión. Para eso, se probaron tres algoritmos distintos con combinaciones de hiperparámetros diferentes al mismo tiempo que buscamos diferentes

combinaciones de *features*. A continuación, detallaremos el plan de acción tomado en cuanto a algoritmos utilizados, métodos de validación y métricas de performance.

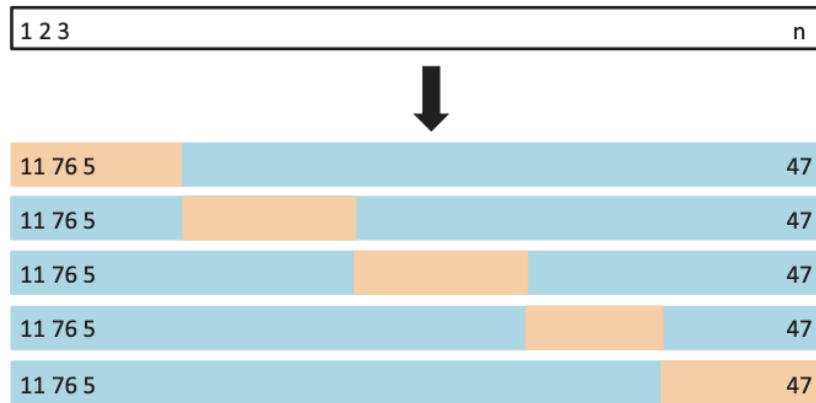
3.3.1 Métodos de validación y métricas de performance

Uno de los componentes más importantes y que buscamos sea consistente entre los distintos algoritmos es el método de validación de los datos. En tal sentido, realizamos una división de los datos entre entrenamiento y testeo manteniendo un 80% en el primero y un 20% en el segundo: esto es, 1668 vendedores para entrenar y 417 para testear. Adicionalmente, sobre el dataset de entrenamiento realizamos un validación cruzada de 5 capas (James, et. al. 2013). Para ello, se divide al dataset en 5 grupos diferentes seleccionando uno como conjunto de validación y entrenando con los otros cuatro. Seguido de esto tomamos el siguiente grupo y realizamos el mismo proceso. Para cada una de las etapas calculamos como métrica de performance la raíz del error medio cuadrático y promediamos su valor como métrica final de entrenamiento:

$$CV = \sum_{i=1}^{k=5} RMSE_i \quad \text{donde } RMSE = \sqrt{\frac{\sum_{t=1}^T (y^{\wedge} - y)^2}{T}}$$

El proceso que acabamos de detallar se ilustra en la Imagen 4. La finalidad de todo este procedimiento es reducir el ruido que se puede incorporar al entrenar el modelo debido a un sobre ajuste sobre los datos de entrenamiento y testeo (Alpaydin, 2014).

Imagen 4. Ejemplificación del proceso de validación cruzada de 5 grupos.



Fuente: An Introduction to Statistical Learning: With Applications in R (Gareth, et. al. 2013)

3.3.2 Modelos utilizados

El primer algoritmo utilizado para resolver el problema fue un modelo lineal simple de regresión utilizando el método de *shrinkage Ridge Regression* (Gareth, et. al. 2013). Este método busca minimizar la función de pérdida definida como la suma de los errores al cuadrado (SEC) más un término adicional definido a través de un parámetro λ que penaliza cuando los coeficientes son muy cercanos a cero.

$$L = SEC + \lambda \sum_{j=1}^p \beta_j^2$$

Cuanto más cercano a 1 es λ , mayor el grado de penalización por coeficientes cercanos a cero y en tal caso nos encontramos en un modelo LASSO. En consecuencia, el conjunto de parámetros beta que produce la regresión será diferente al obtenido por una regresión lineal tradicional. En este trabajo testeamos distintos valores del parámetro λ para obtener el mejor fit del modelo.

El segundo algoritmo que empleamos fue una *Decision Tree Regression*. En términos generales, el proceso consiste en realizar diversos cortes sobre el espacio de predicción y predecir el mismo valor para cada observación perteneciente a determinada región. Es un algoritmo recursivo de tipo *top-down* y *greedy*, en el sentido que realiza una partición inicial

sobre todas las observaciones y cada nuevo paso o partición del árbol la construye buscando la minimización de la suma de errores al cuadrado para esa partición y no buscando la mejor minimización a nivel global (Gareth et. al, 2013). Ahora bien, al momento de realizar cada corte debemos establecer un límite para evitar hacer un *overfit* de los datos. Para ello, determinamos ciertos hiperparámetros que permiten regular la profundidad del árbol. A modo de resumen y considerando lo recién mencionado, el algoritmo funciona de la siguiente manera:

1. Se toma el total del espacio de predicción y una de las variables X_j haciendo una partición buscando la máxima minimización de la suma de los errores al cuadrado posible.
2. Una vez realizado el corte, se repite el proceso de corte utilizando otra variable X_k hasta que la cantidad de observaciones del nodo tiene menos cantidad que la establecida en un hiperparámetro
3. Utilizamos un hiperparámetro α en cada paso que se agrega a la función de costo con fin de evitar tener demasiados nodos terminales y árboles muy profundos que conduzcan a un *overfit* de los datos.

Uno de los problemas del árbol de decisión es la falta de robustez y precisión de la predicción en testeo. Este problema se desarrollará más adelante cuando se presenten los resultados obtenidos con el mismo.

El tercer y último modelo propuesto es un algoritmo de entrenamiento aditivo: *Ext Gradient Boosting Machine*. Conceptualmente, el procedimiento consiste en realizar un ajuste recursivo utilizando diversos árboles de decisión pequeños y actualizando los residuos en cada paso. El modelo final de la predicción será una agregación de cada uno de los árboles ajustados en cada paso. Siendo más específicos, se establece $f_{\text{pred}}(x) = 0$ y $\text{error}_i = y_i$ para todas las observaciones de entrenamiento y tres hiperparámetros B (cantidad de árboles), λ (tasa a la que aprende el algoritmo) y d (cantidad de particiones que se realizarán en cada árbol). Luego para todos los árboles de 1 a B se realizan los siguientes pasos:

1. Se ajusta un árbol con d cortes sobre los datos y se actualiza $f_{\text{pred}}(x)$ siendo ahora $f_{\text{pred}}(x) = f_{\text{pred}}(x) + \lambda f_b(x)$

2. Se actualizan los residuos $error_i = error_i - \lambda f_b(x)$

La definición de los hiperparámetros B , λ y d pueden obtenerse de diferentes maneras. Por un lado, existe el método *grid search* que consiste en crear una matriz con diferentes combinaciones de los mismos y barrer las diferentes posibilidades. Por nuestra parte, decidimos realizar el ajuste de hiperparámetros a través del método *random search*. Este procedimiento establece los valores de los hiperparámetros de manera aleatoria y se queda con aquella combinación que produce el mejor *fit*. Decidimos utilizar esta metodología siguiendo el trabajo de Bergstra y Bengio (2012) quienes muestran teórica y empíricamente *random search* es más eficiente que *grid search* a la hora de encontrar los mejores hiperparámetros.

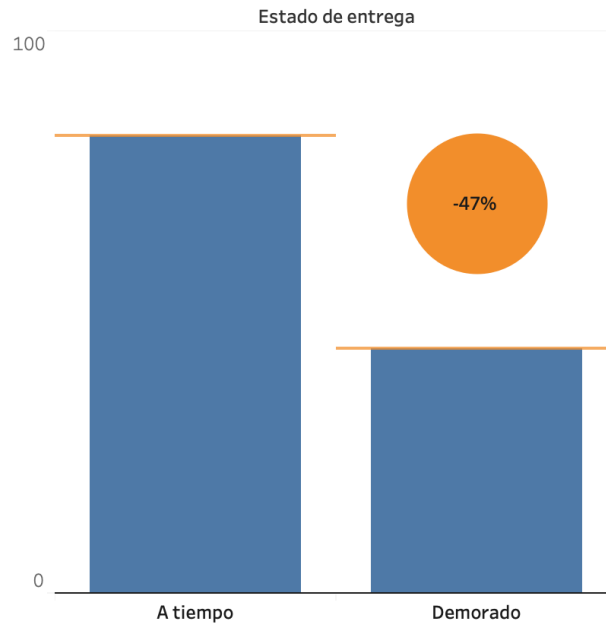
4. Exploración de los datos

A continuación, se profundizará en los datos obtenidos tanto de las encuestas como de las transacciones exploradas. Daremos un detalle de las variables a utilizar explicando el racional detrás de la inclusión de cada una en el modelo. Dividimos la sección en dos partes: una donde exponemos los datos asociados a los envíos y otra donde exploramos el efecto de las cancelaciones y de las diferentes categorías en el NPS. Se fracciona en estas dos secciones debido a que la etapa de envío de un paquete es aquella que más influencia tiene en la experiencia de compra y recepción de un paquete.

4.1. Los datos del envío

Una de las hipótesis iniciales sobre el modelo es que entregar un paquete fuera de plazo genera mala experiencia. Asimismo, si los plazos que el comprador debe esperar por su paquete son más largos, el NPS va a ser peor. Ambas hipótesis surgen de un análisis gráfico del total de las encuestas analizadas. En la Figura 2 se observa el NPS promedio del total de encuestas analizadas abierto por el estado de la entrega de la compra. Para el caso de las compras que fueron entregadas con demora el NPS cae un 47% respecto de aquellas compras que fueron entregadas a tiempo. En la Figura 2, para mantener la confidencialidad de los datos, eliminamos el eje que muestra el número exacto de NPS dejando sólo el diferencial según el estado de la compra.

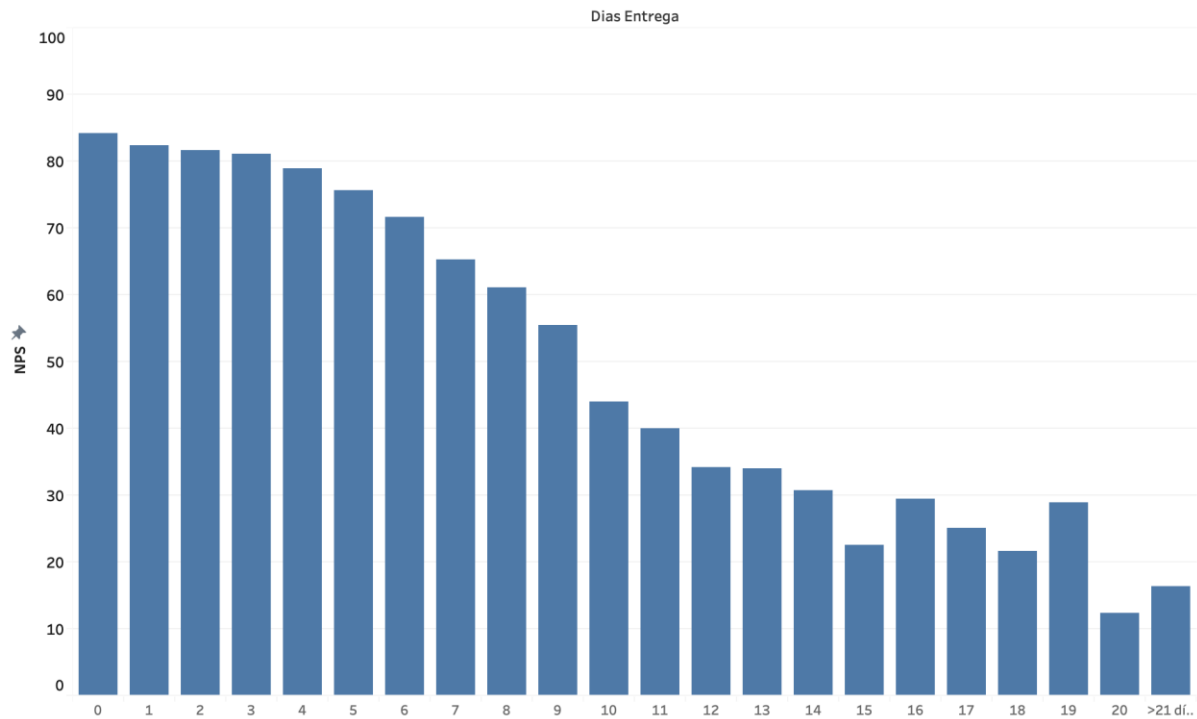
Figura 2: NPS según la demora en la entrega de la compra.



Fuente: elaboración propia (Tableau) en base a encuestas analizadas

Por su parte, la segunda hipótesis que busca reflejar la espera del usuario por su producto. En la Figura 3 se expone que a medida que los días de entrega del producto crecen, el puntaje de NPS tiende a caer. Las caídas del NPS se acentúan cuando la espera supera los 5 o 10 días, siendo del 2% -respecto al día previo- durante los primeros 5 días, en torno a un promedio del 6% hasta el día 10. Luego, el comportamiento es más errático donde podemos ver caídas que pueden llegar hasta el 20% pero a la vez leves subidas de NPS. Encontramos dos hipótesis que podrían explicar este comportamiento. En primer lugar, las promesas de entrega mayores generan una peor experiencia en los usuarios porque deben esperar más tiempo para obtener su producto. En segundo lugar, un día adicional de espera en plazos largos no afecta la satisfacción del usuario dado que sus expectativas de entrega están fijadas en plazos mayores.

Figura 3: NPS según los días de entrega de la compra.



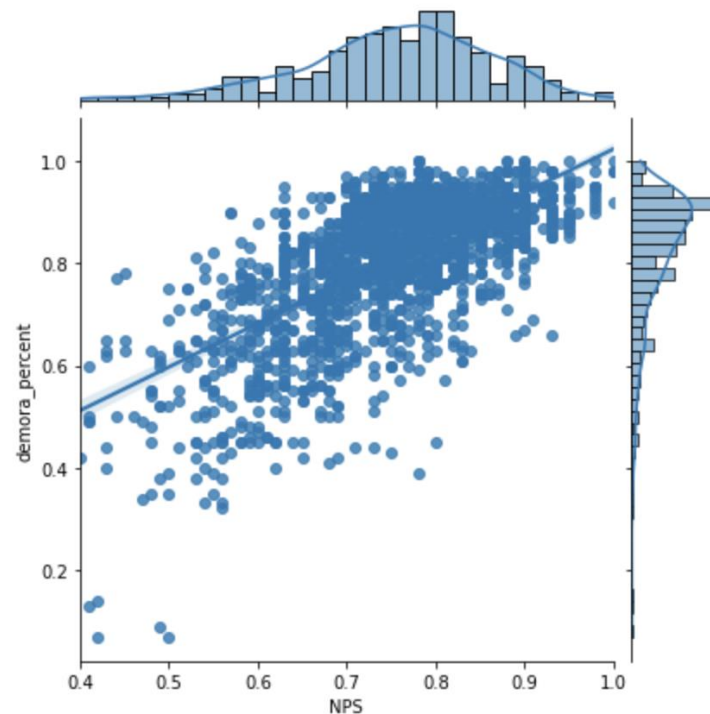
Fuente: elaboración propia (Tableau) en base a encuestas analizada

Dos derivaciones que surgen del análisis de las figuras 2 y 3 es que, a primera vista, los plazos cortos tienden a cumplirse mientras que promesas más largas tienden a performar de una peor manera. Además, otra posible explicación al comportamiento del NPS en ambas figuras es que si se fijan plazos de entrega cortos para los usuarios (hasta 10 días), la expectativa es el cumplimiento de la promesa, por lo tanto es importante llegar a tiempo. Ahora bien, en plazos más largos el incumplimiento pareciera no afectar tanto la experiencia debido a que no hay una expectativa sólida de cumplimiento. En este sentido, es preferible fijar expectativas de un plazo más largo de entrega, y cumplir con la misma a fijar plazos cortos y entregar demorado.

A pesar de que se sostiene que los días de demora y entrega afectan negativamente a la experiencia, los vendedores pueden tener contar con pocas ventas de este tipo en su historial y en el promedio mantener elevados niveles de NPS. La primera aproximación tomada para validar estas hipótesis son métodos gráficos y regresiones lineales. En la figura 4 se observa que el porcentaje de ventas entregadas a tiempo (`demora_percent`) contra el NPS del vendedor. Siguiendo esta gráfica entre mayor es el porcentaje de envíos entregados a tiempo, mayor es el

NPS. Sin embargo, vemos que la correlación lineal entre las variables es leve: 0.31⁸. Sin embargo, lo consideramos como uno de los features a incorporar en el modelo dado que buscaremos relaciones más complejas que las lineales.

Figura 4: Distribución conjunta del NPS vs el porcentaje de entregas a tiempo



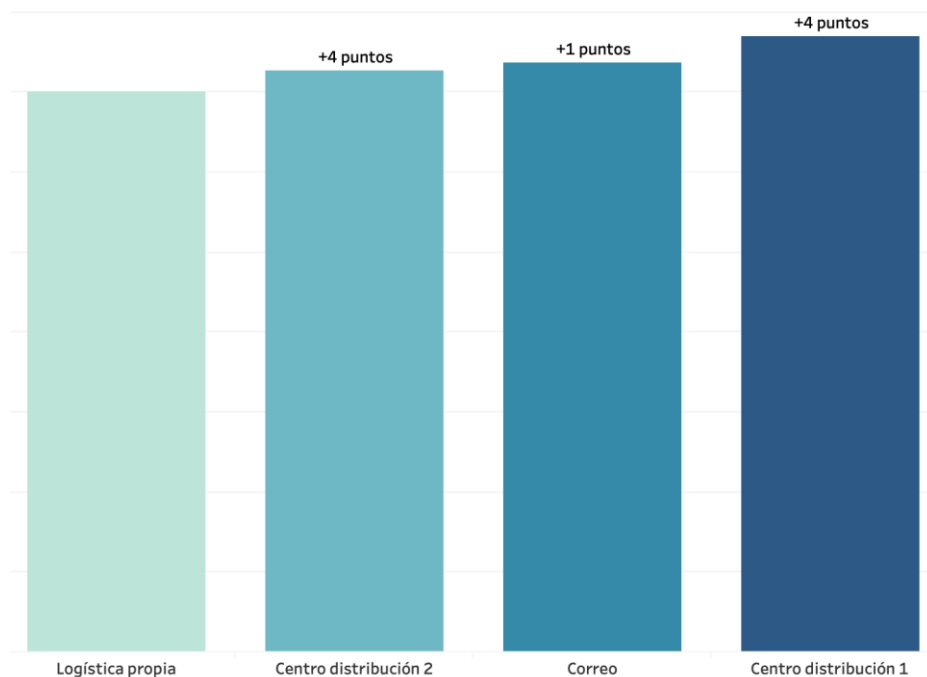
Fuente: elaboración propia en base a los vendedores analizados

El tercer plano relacionado al envío del producto es el método utilizado para el mismo. Una de las variables mencionadas en el Capítulo 2 es MET_ENV que captura el hecho de que el paquete viaje por el correo o servicio postal tradicional del país, si sale de algún centro de distribución de la plataforma o si es enviado por logística propia. La compra puede llegar más rápido a manos del comprador si el paquete se encuentra en uno de los centros de distribución de la plataforma que si el vendedor debe acercarse a un punto de correo y despachar el producto. Además de la velocidad de entrega, que el paquete se encuentre en alguno de los centros de distribución permite mantener mayor control en el almacenaje y así evitar posibles deterioros

⁸ Se realizó la limpieza de 3 observaciones que parecieran ser outliers en la figura 4 y la correlación disminuyó levemente a 0.26

en el proceso de embalaje y despacho. Es por estos motivos que esperamos que el NPS de los centros de distribución sea superior a la logística propia o entregas por correo. Sin embargo, es necesario hacer una distinción sobre el tipo de centro de distribución: puede ser propio de la plataforma (Centro distribución 1) o un hub tercerizado (Centro distribución 2). Como puede observarse en la Figura 5, el NPS de los paquetes enviados por logística propia es menor al resto de los otros métodos de envío. Por su parte, aquellos que salen del centro de distribución 2 tienen 4 puntos más de NPS y los del centro de distribución 1, 9 puntos más, ambos comparados contra la logística propia. Por lo tanto, consideramos que el método de envío es una feature importante para poder detectar diferencias de NPS y serán incorporadas en las combinaciones de variables del modelo.

Figura 5: NPS según el método de envío de la compra.



Fuente: elaboración propia (Tableau) en base a las encuestas analizadas

4.2. Cancelaciones, categorías y NPS

Otro de los puntos que pueden afectar la experiencia de los usuarios es el hecho de que la compra haya sido cancelada. La anulación de una compra puede realizarse por la falta de cumplimiento en la promesa de entrega, por un arrepentimiento o error del comprador a la hora

de realizar la transacción, por situaciones particulares del vendedor como quiebre de stock o problemas con el despacho. Todas estas situaciones generan que la transacción no sea exitosa y conducen fricciones entre el comprador y el vendedor. Para exponer este efecto contamos con el Figura 6 en el que se observa el NPS de los vendedores analizados según el porcentaje de cancelaciones de las ventas analizadas. Aquellos con cancelaciones menores al 3% presentan NPS por encima de los 75 puntos, mientras que en el otro extremo, los porcentajes de cancelaciones superiores al 10% no alcanzan los 70 puntos de NPS.

Figura 6: NPS según el porcentaje de cancelaciones del vendedor

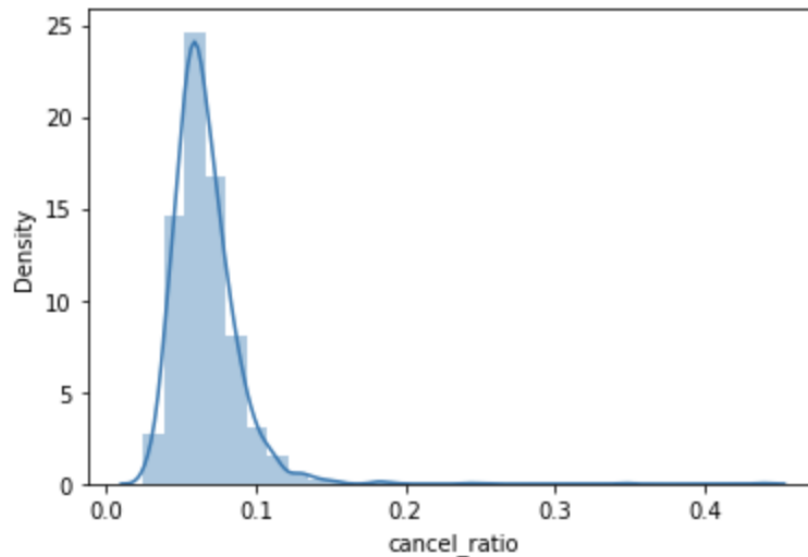


Fuente: elaboración propia (Tableau) en base a los vendedores analizados

Como complemento al figura 6, exponemos la distribución de la tasa de cancelaciones de los vendedores en el figura 7. De aquí se desprende que en la base que utilizaremos para hacer el fit del modelo, el porcentaje de cancelaciones está concentrado en valores cercanos al 7%. Sin embargo, tenemos algunas observaciones que claramente son valores *outliers*, con ratios de cancelación que superan el 40%. En otras palabras y según esta información, casi una de cada dos transacciones de este vendedor con comportamiento atípico termina sin ser exitosa. En términos del modelo que buscamos generar, consideramos que tener una observación de

este estilo genera demasiado ruido a la base y decidimos eliminar esta observación a la hora de hacer el entrenamiento.

Figura 7: Distribución del ratio de cancelaciones de los vendedores analizados

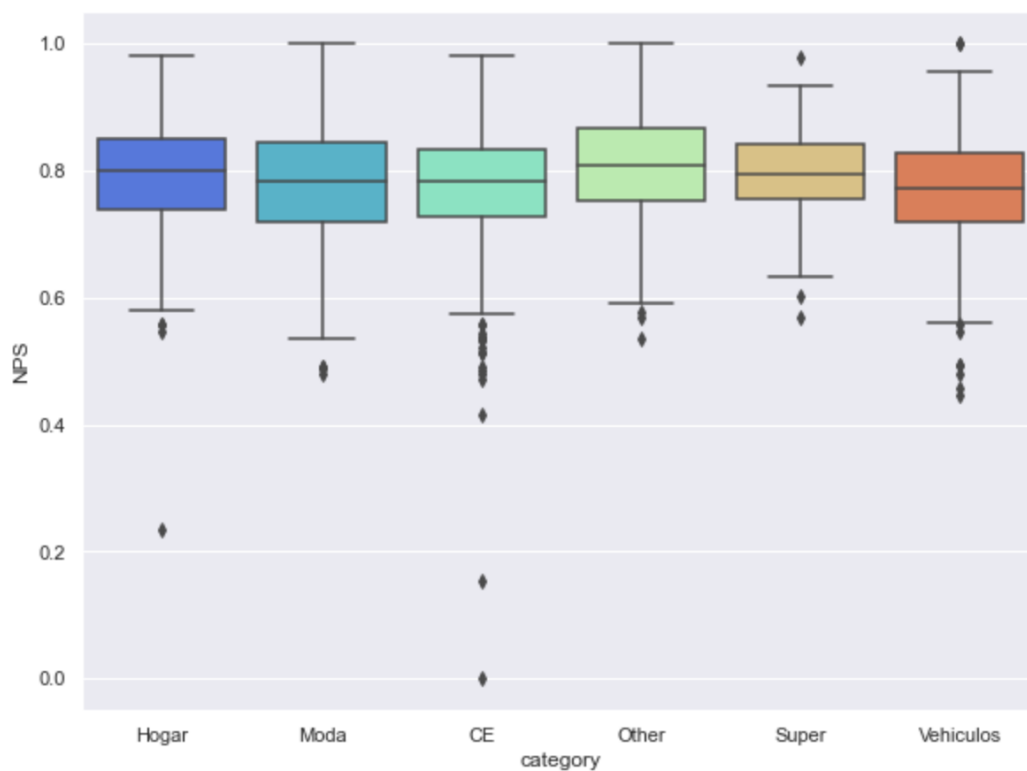


Fuente: elaboración propia en base a los vendedores analizados

De la misma forma en que la calidad del producto puede verse afectada por su tratamiento durante la etapa de envío y con ello la experiencia, el tipo de producto o categoría de venta también puede alterar el valor del NPS. Una hipótesis que surge de esto es que existen ciertos tipos de producto de consumo suntuario, como pueden ser los celulares o consolas, que se tienden a comprar sin demasiado análisis por parte de los compradores con el objetivo de aprovechar cierta oferta. Sin embargo, las condiciones que se proponen en la publicación no son las óptimas: productos usados, de peor calidad o con promesas de entregas más largas. Con el objetivo de ganar en cuanto a precio, los compradores están resignando otros aspectos de la experiencia de compra y esto puede afectar el valor de NPS. En nuestra base, vemos el ejemplo en la categoría Consumer Electronics que mantiene un valor de NPS promedio 3 puntos menor a otras categorías. Adicionalmente, tiene una mayor cantidad de valores *outliers* que empujan este promedio hacia abajo. La categoría que mantiene mejor performance es la de Hogar (primera distribución desde la izquierda en la Figura 8). A pesar de que Consumer Electronics

en conjunto con Accesorios para vehículos estén levemente por debajo del resto, el NPS entre categorías se mantiene estable entre los 70 y 80 puntos. A continuación exhibimos las distribuciones del NPS escalado utilizando MinMax Scaler para las 6 posibles categorías de venta: Accesorios para vehículos, Moda, Consumer Electronics, Artículos para el Hogar, Otras Categorías y Artículos de supermercado.

Figura 8: Distribución del NPS escalado por categoría

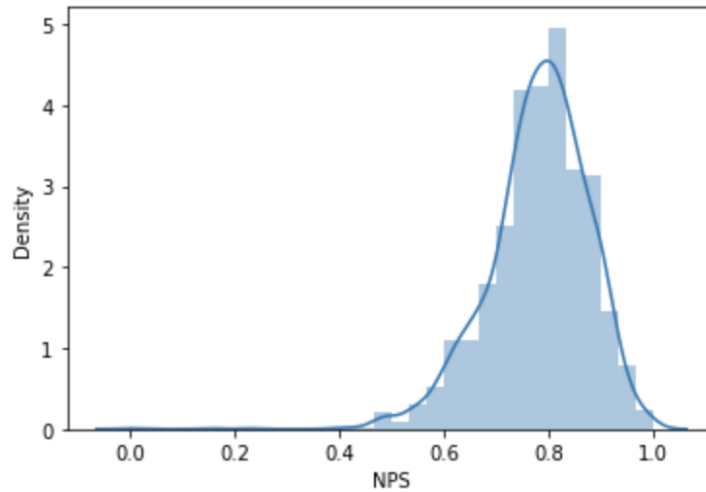


Fuente: elaboración propia en base a los vendedores analizados

Por último, cabe mencionar un problema al que nos enfrentamos al intentar predecir NPS. Este consiste en la baja cantidad y variabilidad de los datos obtenidos para entrenar el modelo. Contamos con 2085 datos de NPS donde el 50% de los valores centrales de la

distribución se encuentran entre los 70 y 82 puntos de NPS. A continuación exhibimos el punto en el figura 4 donde puede observarse la distribución de los datos de NPS escalados⁹.

Figura 9: Distribución del NPS escalado utilizando min max scaler



Fuente: elaboración propia en base a los vendedores analizados

⁹ Escalamos los datos de NPS para mantener su confidencialidad

5. Resultados

5.1. El modelo base

Actualmente en la compañía se establecen un objetivo de NPS de los compradores que se espera que al menos los grandes vendedores cumplan a fin de cada año. El fin último de fijar un objetivo es crear y mantener altos niveles de experiencia y satisfacción a la hora de transaccionar en la plataforma. Cuando los asesores comerciales de las empresas vendedoras de la plataforma detectan desviaciones de estos objetivos pueden tomar distintas acciones para mejorar ese KPI.

Como modelo base de comparación de los algoritmos mencionados en la sección anterior, utilizaremos la metodología de cálculo de estos objetivos. El valor de NPS predicho será entonces un único valor fijo para todos los vendedores analizados y se crea según el siguiente procedimiento:

1. Tomamos todas las encuestas de los 6 meses previos al momento del análisis. En este caso, a Octubre 2020.
2. Computamos el NPS promedio de cada una de las categorías de venta para cada mes.
3. Calculamos la cantidad de ventas de cada categoría para el mismo período.
4. Para cada mes, calculamos un NPS_t que consiste en el promedio de los NPS de las categorías ponderado por la cantidad de transacciones de esa categoría en ese mes.
5. Por último, tomamos el NPS promedio de los NPS_t

$$NPS_{baseline1} = \frac{\sum_{t=1}^6 NPS_t}{6} \quad NPS_{baseline2} = \frac{\sum_{t=1}^4 NPS_t}{4}, \text{ donde } NPS_t \text{ no es máximo ni mínimo}$$

Como alternativa, y para evitar picos donde el NPS puede haber sido un *outlier* tomamos creamos un segundo modelo baseline que consiste en el mismo procedimiento pero quitando el NPS máximo y mínimo los últimos 6 meses. En la tabla 5 se observa el resumen

de la performance esta metodología, en el cual obtuvimos en testeo una raíz del error medio cuadrático (RMSE) de 0.10722 para el baseline 1 y de 0.10724 para el baseline 2 (aquel en el que excluimos outliers). El objetivo de nuestros algoritmos es obtener un mejor RMSE que estos dos modelos.

Tabla 5: resumen modelo baseline

Model	RMSE train	RMSE test
Modelo base 1	0.10464	0.10722
Modelo base 2 (suavizado)	0.10464	0.10724

5.2. Combinaciones de variables

En función de clarificar la exposición de los distintos modelos utilizados, primero detallaremos las cinco combinaciones de variables utilizadas. Para los tres modelos y sus respectivos hiperparámetros realizamos el fit de 5 combinaciones distintas de variable:

1. **Combinación 1 (M1):** Contiene 29 features que agrupamos en los siguientes 16 grupos.
 - 1) Días de entrega
 - 2) Demora percent
 - 3) Cantidad de ítems de compra
 - 4) Cantidad de cancelaciones 30 días previos
 - 5) Cantidad de reclamos 30 días previos
 - 6) Porcentaje de reclamos (claim_ratio)
 - 7) Porcentaje de cancelaciones (cancel_ratio)
 - 8) Porcentaje de usuarios de compra frecuente

- 9) Porcentaje de usuarios de compra regular
 - 10) Porcentaje de usuarios nuevos
 - 11) Porcentaje de compras compradas con efectivo, con tarjeta de crédito, con tarjeta de débito y con dinero digital
 - 12) Categoría de ventas: accesorios, consumer electronics, supermercado, categoría hogar
 - 13) Índice de concentración: promedio (HH_INDEX), desvío estándar del HH_INDEX
 - 14) Ventas promedio y desvío estándar de ventas mensuales
 - 15) Promedio y desvío estándar de ventas diarias
 - 16) Tipos de envíos: porcentaje de envíos con logística propia del vendedor, porcentaje de envíos con correo tradicional, porcentaje de envíos que salen del centro de distribución contratado por la plataforma o propio de la misma.
2. **Combinación 2 (M2):** Contiene 34 features, siendo 29 las pertenecientes a la combinación M1 pero agregando los desvíos estándar de los métodos de envío de los paquetes.
3. **Combinación 3 (M3):** Contiene 28 variables que agrupamos en 11 grupos.
- 1) Días de entrega: el promedio, desvío y mediana
 - 2) Demora percent
 - 3) Cantidad de ítems de compra
 - 4) Cantidad de cancelaciones 30 días previos
 - 5) Cantidad de reclamos 30 días previos
 - 6) Mediana de los reclamos (claim_ratio)
 - 7) Mediana de las cancelaciones

- 8) Tipos de compradores: porcentaje de usuarios de compra frecuente, porcentaje de usuarios de compra regular, porcentaje de usuarios nuevos
 - 9) Medios de pago: porcentaje de compras compradas con efectivo, con tarjeta de crédito, con tarjeta de débito y con dinero digital
 - 10) Categoría accesorios, categoría consumer electronics, categoría supermercado, categoría hogar
 - 11) Índice de concentración: promedio (HH_INDEX) y desvío estandar
4. **Combinación 4 (M4):** Contiene 27 variables modificando aquellas referidas a la categoría de venta del vendedor. En lugar de utilizar el porcentaje de ventas de cada una de las categorías, generamos variables dummy que toma valor 1 si el la categoría analizada es aquella donde el vendedor vende la mayoría de sus productos. Por otra parte, eliminamos las variables de método de envío.

5.3. Performance de los modelos

Considerando las diferentes combinaciones de variables, ajustamos en primer lugar un modelo Ridge Regression. La combinación de variables que resultó con mejor performance, utilizando como medida de la misma el RMSE, fue la M1. Para cada una de las pruebas con los distintos alphas, la combinación M1 demostró tener un menor RMSE en training (y testeo) siendo 0.0887 el menor alcanzado. De aquí se desprende la inferencia que el modelo más básico en cuanto a combinación de variables ya contiene la información que ayuda a predecir el valor de NPS de la mejor manera y que agregar nuevos features para la cantidad de observaciones existentes no agrega valor. Por otra parte, al comparar los RMSE podemos observar que el parámetro alpha no modifica sustancialmente los valores de train entre las iteraciones. Los valores 0.4 y 0.5 generaron la mejor combinación posible en train, pero cuando vemos los valores de test, alpha = 0.5 ha generado el menor RMSE.

Tabla 6: resumen modelo Ridge Regression

Combinación de variables	Parámetro alpha	RMSE train	RMSE test
M1	0.4	0.0887	0.0891
M2	0.4	0.0888	0.0889
M3	0.4	0.0912	0.0930
M4	0.4	0.0904	0.0919
M1	0.5	0.0887	0.0892
M2	0.5	0.0888	0.0889
M3	0.5	0.0912	0.093
M4	0.5	0.0904	0.0919
M1	0.6	0.0887	0.0893
M2	0.6	0.0888	0.089
M3	0.6	0.0913	0.093
M4	0.6	0.0905	0.092
M1	0.7	0.0887	0.0893
M2	0.7	0.0888	0.089
M3	0.7	0.0914	0.0931
M4	0.7	0.0905	0.0921

Por otro lado, vemos que la performance del árbol de decisión no supera a la regresión realizada mediante Ridge. Hemos utilizado como parámetros default que el criterio para hacer el corte sea la reducción del MSE (*Mean Squared Error*), no se ajustó la cantidad de variables para realizar cada corte y que el mínimo de hojas para realizar un *split* sea 1. Alteramos la profundidad del árbol con el objetivo de evitar un sobre ajuste en los datos de entrenamiento. Como puede observarse en la Tabla 6, el modelo con el mejor ajuste es aquel que utiliza una

profundidad de árboles limitados en 3 y la combinación de variables M3, es decir, aquella que incorpora las medianas de las cancelaciones y reclamos generados por los usuarios. En este caso, el RMSE de entrenamientos obtenido es 0.0933 superando en apenas 1% a la segunda mejor combinación de variables y límite de 3 en la profundidad de los árboles. Sin embargo, en comparación con el modelo propuesto más arriba, el RMSE del mejor ajuste de la regresión por árboles de decisión es 2% superior al peor ajuste de la regresión del modelo Ridge. Es por este motivo que descartamos al árbol de decisión como modelo para resolver nuestro problema.

Tabla 7: resumen modelo Decision Tree Regressor

Combinación de variables	Máxima profundidad	RMSE train	RMSE test
M1	2	0.0961	0.1005
M2	2	0.0961	0.1005
M3	2	0.0949	0.0984
M4	2	0.0961	0.1005
M1	3	0.0945	0.0955
M2	3	0.0944	0.0955
M3	3	0.0933	0.0984
M4	3	0.0946	0.0955
M1	4	0.0941	0.0943
M2	4	0.0944	0.0944
M3	4	0.0988	0.0972
M4	4	0.094	0.0943

En último lugar, utilizamos el algoritmo de entrenamiento aditivo Gradient Boosting Machine. En este caso, la estrategia consistió en crear una función en Python que aplique *Random Search* para poder encontrar los mejores hiperparámetros del modelo. Dicha función hace un fit de 50 estimadores con diferentes hiperparámetros seleccionados aleatoriamente de una grilla de posibilidades. Un punto importante a mencionar, es que la selección de

hiperparámetros mediante esta metodología genera distintos resultados cada vez que se vuelve a ejecutar el código. Por lo tanto, podemos tener mejores o peores resultados por el simple hecho de volver a correr el Notebook de Jupyter donde hemos realizado el estudio. La elección de los mismos se realiza según el mejor RMSE de *cross validation* con 3 capas y el resumen de los resultados con aquellos hiperparámetros óptimos se exhiben en la Tabla 8. La combinación de variables que condujo a un menor RMSE de cross validation fue la M4 con un valor de 0.0973. Para conseguir este resultado, la tasa de aprendizaje del algoritmo λ toma un valor de 0.28, mientras que la cantidad de particiones que realiza cada árbol fue de 15. Adicionalmente, la fracción de features que utilizó el algoritmo para poder entrenar cada uno de los árboles fue de 1.

Tabla 8: resumen modelo Gradient Boosting Machine

Parámetros	M1	M2	M3	M4
λ	0.48	0.48	0.45	0.28
<i>n estimators</i>	6	6	6	15
<i>colsample</i>	0.78	0.58	0.69	1.0
<i>num leaves</i>	28	75	140	60
<i>sub sample</i>	0.58	0.64	0.63	0.62
<i>reg alpha</i>	0.51	0.71	0.55	0.59
RMSE	0.0993	0.0989	0.1014	0.0973

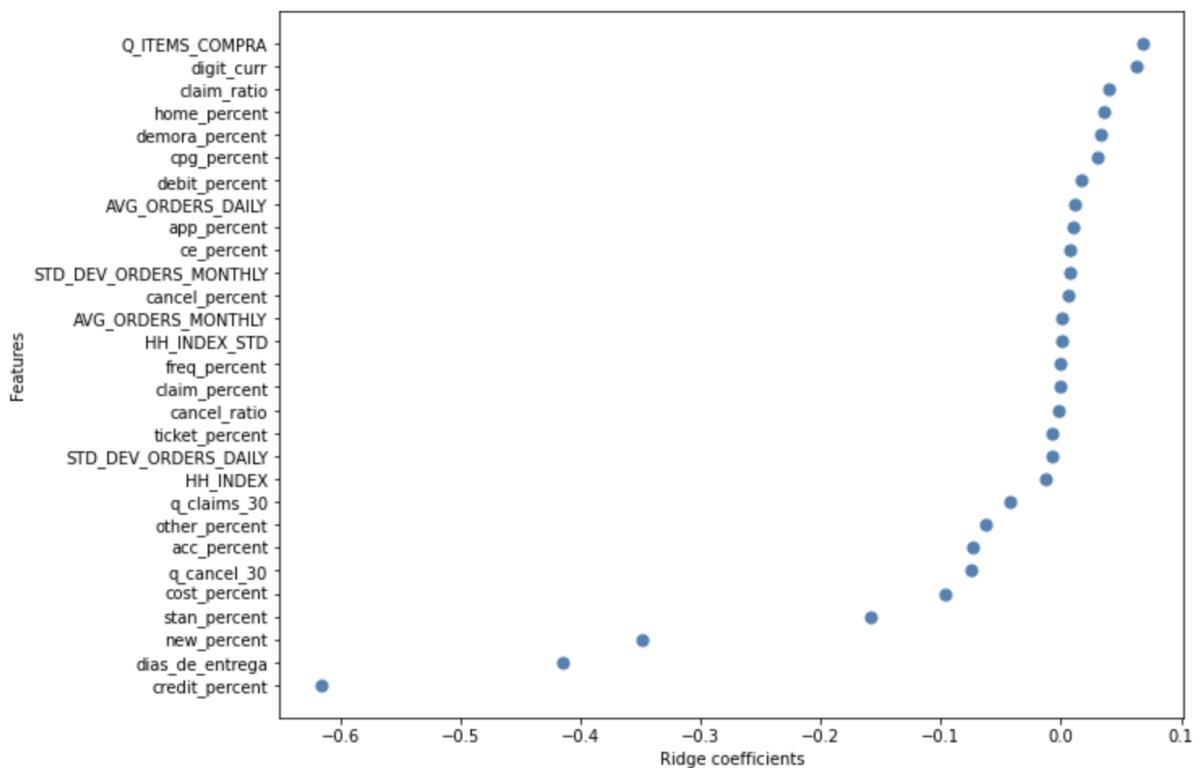
En resumen, de los tres tipos de algoritmos utilizados, aquel que produjo la mejor performance es Ridge Regression. El mejor de los modelos de Ridge generó un RMSE 4.9% menor al mejor de los árboles de decisión y 9.2% respecto de Gradient Boosting Machine. Una posible explicación a este efecto es que dada la limitada cantidad de observaciones (2085 en total con un 20% que fueron destinadas a testing), los algoritmos más complejos no pueden encontrar relaciones demasiado complejas en los datos. Por otra parte, es importante notar que el modelo propuesto de Ridge Regression es superador al modelo *Baseline* en un 15.2%. Es decir, utilizando esta metodología podemos obtener resultados más precisos a la hora de predecir el NPS de los compradores sin la necesidad de contar con las encuestas suficientes.

De esta manera, se podrá disminuir la dependencia de las encuestas cuando se busque mantener un control sobre las posibles desviaciones de NPS de los vendedores de la plataforma.

5.4. Evaluación del modelo

Como se muestra en la sección anterior, el modelo con mejor performance fue el de Ridge Regression con la combinación M1 de variables. Profundizando en lo coeficientes que resultan del ajuste del modelo Ridge se destaca la importancia de las variables asociadas tanto a la compra como al comprador. La Figura 10 muestra la importancia de las variables medida según la magnitud de los coeficientes ajustados por el modelo mejor Ridge.

Figura 10: Importancia de las variables según los coeficientes de Ridge



Fuente: elaboración propia en base a los resultados del modelo Ridge con variables M1

Uno de los resultados más interesantes es que la *feature* de mayor magnitud que encontró el modelo es el porcentaje de ventas con tarjeta de crédito. La dirección negativa de este coeficiente nos dice que cuanto mayor sea el porcentaje de ventas con tarjeta de crédito,

peor será el NPS del vendedor. Una posible explicación de este fenómeno consiste en que los objetos que se compran con crédito son más costosos que aquellos que se compran de contado y esto es un potencial factor de peor experiencia para el comprador. Un elemento adicional de las compras con crédito es el pago de intereses para poder financiar la compra que lleva a que el precio final pagado por el comprador sea más elevado. Una alternativa que ofrece la plataforma a los vendedores como solución a este problema es la posibilidad de ofrecer cuotas sin interés en las compras con tarjeta de crédito haciéndose cargo del costo de esa financiación.

Como se mencionó en el Capítulo 4, uno de los factores que afecta negativamente a la experiencia es la cantidad de días que se debe esperar por el paquete. En la Figura 10 se exhibe que la segunda *feature* con mayor importancia es la cantidad de días de entrega de la compra que tiene signo negativo. Adicionalmente, se muestra como *demora_percent* que captura el porcentaje de envíos a tiempo, tiene un signo positivo pero de importancia menor. En este sentido, es importante para los vendedores poder ofrecer una menor cantidad de días de espera para sus compradores a fin de mejorar la experiencia de compra.

Por último, es interesante mencionar las características del comprador asociadas a la compra. La tercera y cuarta *feature* de mayor relevancia y que impactan negativamente en el NPS son el porcentaje de usuarios nuevos y el porcentaje de usuarios con compras estándar. Cuantos más compradores sin experiencia sean clientes del vendedor menor será la satisfacción de los mismos. Esto da cuenta de la importancia de un *onboarding* o primeras experiencias satisfactorias dentro de la plataforma. Los vendedores pueden ayudar a que los nuevos usuarios se sientan seguros utilizando la plataforma siendo transparentes en las condiciones de compra y de devolución de los productos. En paralelo, deben ser claros en la descripción de los productos que venden así como en las imágenes que publican de esos productos. Otra posible acción que pueden tomar los vendedores es invitar a hacer preguntas previas a la compra y responder las mismas con claridad. Finalmente, en el proceso posterior a la compra, los vendedores tienen la posibilidad de mantener informados sobre el despacho y entrega del producto a través de la mensajería que disponibiliza la plataforma. De esta manera, el comprador tiene visibilidad durante todo el proceso de su compra.

6. Discusión

6.1. Resumen de resultados

Para poder resolver el problema de la falta de encuestas a la hora de obtener un valor representativo de NPS, se propuso generar un modelo que nos permita predecir el valor del mismo mediante un conjunto de variables. Se utilizaron diversos algoritmos de Machine Learning: Ridge Regression, Decision Tree Regression y Gradient Boosting Machine. Como método de validación de los datos separamos en datos de entrenamiento y datos de testeo en conjunto con un 5-fold cross validation. Los tres algoritmos fueron puestos a prueba con diferentes combinaciones de variables y comparados contra un modelo base utilizado actualmente por la compañía.

Para el modelo baseline se usó un promedio ponderado de los NPS mensuales considerando los últimos 6 meses previos al momento del análisis. El ponderador utilizado es la cantidad de ventas de cada una de las categorías (moda, electrónica, hogar). Este modelo baseline predice un único NPS para todos los vendedores analizados y el RMSE obtenido en entrenamiento fue de 0.10464. Todos los algoritmos propuestos arrojaron resultados superiores al modelo base.

En primera instancia, el modelo Ridge con la combinación de variables más básica (M1) demostró tener un RMSE 15.2% ($RMSE = 0.0887$) menor al modelo base. El segundo modelo ajustado, Decision Tree Regressor, tuvo una performance 4.9% peor que la primera medida en términos del RMSE. En este caso, la combinación de variables utilizada para alcanzar el error mínimo fue la M3 con un RMSE de 0.0933. Por último, el algoritmo Gradient Boosting Machine produjo el mayor RMSE 0.0993 que, si bien es superior al modelo base, no logra equipararse a los dos modelos previamente ajustados. En este sentido, la recomendación es utilizar el modelo propuesto con el algoritmo Ridge Regression utilizando la combinación de variables M1 que es más preciso que el modelo actual utilizado por la compañía para determinar un valor de NPS.

6.2. Aplicaciones prácticas

La implementación de la solución propuesta permitirá en el futuro predecir el NPS de un vendedor utilizando la información transaccional e histórica del mismo sin la necesidad de contar con las encuestas suficientes. Diariamente se cuenta con información sobre la cantidad de operaciones de los vendedores, sus reclamos, cancelaciones, costos de envío y todos aquellos datos necesarios para construir las features del modelo propuesto. Es por esto que nuestra solución puede funcionar en un entorno productivo consumiendo la información, generando el ETL, transformando las variables y ajustando el modelo para predecir NPS. A esto se suma la importancia de sintetizar en un único KPI comprensible como es el NPS que radica en poder contar con una mayor velocidad de reacción ante posibles desviaciones de experiencia que se espera que brinden los vendedores. Nuestro modelo permite contar diariamente con un valor de NPS para cada vendedor a pesar de no tener encuestas para poder calcularlo. Es así como un asesor comercial de la plataforma puede dar seguimiento a la métrica de su cartera asesorada, entender cuáles de ellos se están desviando de la performance esperada y tomar las acciones que crean pertinentes.

En caso de ponerlo productivo, podemos generar alarmas y *dashboards* que nos alerten que vendedores están desviando de los objetivos propuestos o están dando una mala experiencia a los compradores. En consecuencia, el equipo comercial encargado de aconsejar y gestionar a ese vendedor obtendrá en tiempo real una estimación precisa al valor de NPS y tomar acciones: pueden proponer al vendedor mudar cierto porcentaje de la mercadería a distintos centros de distribución para ganar velocidad en las entregas, alertar sobre problemas en el correo o *carriers* que entregan la mercadería, asegurarse de que la calidad de los productos vendidos sea la prometida o pausar publicaciones y penalizarlos si la desviación es demasiado grande.

6.3. Limitaciones

Uno de los puntos de mejora que detectamos del estudio realizado es la falta de observaciones o datos para entrenar el modelo. En el presente trabajo nos enfocamos en los 2085 vendedores más grandes de la plataforma para poder tener a aquellos que reciben asesoramiento comercial por parte de la compañía. Una manera de obtener una mayor cantidad de observaciones es extender el análisis a más vendedores a pesar de que su facturación no sea

tan grande. Dichos vendedores pueden tener un comportamiento similar al estudiado en este proyecto y permitirnos extraer más información de los datos. Una segunda alternativa, es agregar más países al análisis dado que la compañía tiene actividades en más de uno.

Considerando las oportunidades de mejora en términos de *features*, podemos mencionar la incorporación de datos sobre el *backlog* del correo o carrier. Con esto nos referimos a la cantidad de paquetes que el correo debe repartir. A menos que el vendedor utilice su logística propia, no puede responsabilizarse de las demoras generadas en el correo y que pueden derivar en una mala experiencia para el comprador. Incorporar esta información podría ser interesante para poder capturar posibles demoras en la entrega del paquete así como cancelaciones por demora o reclamos realizados.

Una aproximación diferente para resolver el problema y poder contar con más observaciones es intentar predecir si el resultado de una transacción determinada generaría que el comprador de esta sea promotor, detractor o pasivo. En este caso, el problema de regresión que planteamos para predecir el NPS que obtiene el vendedor se transformaría en un problema de clasificación. De esta manera, la pregunta que debemos resolver pasaría a ser: el usuario de una compra determinada, ¿será un promotor, detractor o neutro? Una vez resuelta esta pregunta para cada una de las ventas del vendedor en un día determinado, se realiza el cálculo del KPI y se obtiene el valor de NPS para ese día. Esta aproximación tiene una ventaja adicional a la ganancia en observaciones: la posibilidad de capturar las variables temporales en el momento en que suceden. En este sentido, se evitaría tener que hacer agregaciones de las variables en diferentes momentos del tiempo. Es decir, en lugar de hacer una agregación sobre el promedio de las ventas de los últimos 30 días, podría obtener la foto de la facturación en ese instante del tiempo generando posiblemente mayor precisión en el dato. Dos ejemplos adicionales de variables temporales que sobre el cuál se evitaría tomar una decisión sobre qué valor tomar son las cancelaciones y reclamos. Al momento de la transacción podríamos saber cuántas cancelaciones tuvo el vendedor en el último tiempo y cuál es la tendencia al momento de la compra y la predicción. Por otro lado, podemos saber cuantos reclamos tiene abiertos en ese momento, así como cuál es la tasa a la que viene resolviendo los problemas con sus compradores.

6.4. Conclusión

El objetivo del presente trabajo fue encontrar una metodología que permita predecir el *NPS de los compradores* para aquellos vendedores asesorados para los que la plataforma no tiene encuestas suficientes. El NPS es uno de los principales KPIs utilizados por la compañía para evaluar los niveles de satisfacción y experiencia de los clientes. El *NPS de los compradores* nos da indicios sobre la experiencia de compra en la plataforma y sobre la performance de los vendedores a la hora de transaccionar en el mismo. Podemos calcular el valor de NPS de un vendedor de la plataforma mirando las encuestas que contestan sus compradores y de esta manera identificar si es necesario tomar medidas para garantizar que la experiencia transaccional sea la óptima. Sin embargo, para tener un valor representativo de NPS que permita tomar decisiones sobre el mismo, es necesario contar con una mínima base de respuestas que no siempre es factible de conseguir. El origen de la falta de encuestas es la baja tasa de respuesta de encuestas que mantiene el envío de las mismas por correo electrónico. Adicionalmente, no se puede enviar las encuestas de manera repetida a los clientes para que los mismos no consideren a las comunicaciones de la empresa como mero *spam*.

Para poder resolver este problema se propuso la utilización de diferentes algoritmos de Machine Learning. Previamente, debió ser generado un dataset que nos permita contar con las features necesarias. Se enviaron 160.000 más encuestas y se usaron datos desde 2019 hasta Septiembre 2020. En total, se contaba con 2085 observaciones de diferentes vendedores. Al mismo tiempo se creó un dataset que agrupaba variables con información asociada a los compradores, a los vendedores y a la transacción misma. Con esta información cruda se crearon diferentes *features* que sirvieron de *input* para el modelo.

Finalmente, se ajustaron tres algoritmos distintos con diferentes combinaciones de variables para predecir NPS utilizando como método de validación *k-fold cross validation*. El algoritmo que a primera instancia parecía más simple, Ridge Regression, resultó aquel con mejor performance superando a *Regression Tree* y al *Gradient Boosting Machine*. Además, demostró ser 15,2% mejor que el modelo base establecido: un promedio de los NPS de los últimos 6 meses ponderado por las ventas de cada una de las categorías de artículos disponibles en la plataforma.

Si bien el modelo propuesto tiene oportunidades como son la incorporación de nuevas variables y la ampliación de la cantidad de observaciones, ha demostrado tener una performance superadora a la solución actual utilizada por la compañía. Adicionalmente,

permite contar diariamente con un valor de NPS para cada vendedor sin la necesidad de contar con encuestas suficientes. Los empleados de la empresa encargados de asesorar a los vendedores de la plataforma pueden contar con un KPI que les permita detectar desviaciones en la experiencia ofrecida por los vendedores y ayudar a garantizar la mayor satisfacción de los compradores.

7. Bibliografía

- Aburrous, Maher, Hossain, Alamgir, Dahal, Keshav and Thabtah, Fadi (2010) Intelligent phishing detection system for e-banking using fuzzy data mining. *Journal of Expert Systems with Applications*, 37 (12). pp. 7913-7921. ISSN 0957-4174
- Almquist, E., Wynn G., Cleghorn J. & Sherer L. (2018, Octubre 10). *Delivering What Consumers Really Value*. Bain & Company. <https://www.bain.com/insights/delivering-what-consumers-really-value/>
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).
- Campaign Monitor (2020). *Ultimate Email Marketing Benchmarks for 2021: By Industry and Day*. Campaign Monitor. <https://www.campaignmonitor.com/resources/guides/email-marketing-benchmarks/>
- Cheong, H. J., & Morrison, M. A. (2008). Consumers' reliance on product information and recommendations found in UGC. *Journal of Interactive Advertising*, 8(2), 38-49.
- Du Toit G., Anderson & Hatherall, R., E., (2020, Septiembre 08). *Tools and Trends: Let No Tool Stand Alone*. Bain & Company. <https://www.bain.com/insights/customer-experience-tools-and-trends-2020-let-no-tool-stand-alone>
- Davis, A., & Khazanchi, D. (2008). An empirical study of online word of mouth as a predictor for multi-product category e-commerce sales. *Electronic markets*, 18(2), 130-141.
- Fader, P. (2020). *Customer centricity: Focus on the right customers for strategic advantage*. Wharton digital press.
- Gentile, C., Spiller, N., & Noci, G. (2007). How to sustain the customer experience: An overview of experience components that co-create value with the customer. *European management journal*, 25(5), 395-410.
- Grisaffe, D. B. (2007). Questions about the ultimate question: conceptual considerations in evaluating Reichheld's net promoter score (NPS). *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 20, 36.

- Hagiu, A. (2009). Multi-sided platforms: From microfoundations to design and expansion strategies. *Harvard Business School Strategy Unit Working Paper*, (09-115).
- Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecom churn management. *Expert Systems with Applications*, 31(3), 515-524.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.
- Krol, M. W., de Boer, D., Delnoij, D. M., & Rademakers, J. J. (2015). The Net Promoter Score—an asset to patient experience surveys?. *Health Expectations*, 18(6), 3099-3109.
- Kuo, Y. F., Wu, C. M., & Deng, W. J. (2009). The relationships among service quality, perceived value, customer satisfaction, and post-purchase intention in mobile value-added services. *Computers in human behavior*, 25(4), 887-896.
- Michelli, J. A. (2007). The Starbucks experience: 5 principles for turning the ordinary into the extraordinary. *Esensi*
- Nielsen, J., Molich, R., Snyder, C., & Farrell, S. (2000). E-commerce user experience. Nielsen Norman Group.
- Pei-Yu Chen, & Hitt, L. M. (2002). Measuring Switching Costs and the Determinants of Customer Retention in Internet-Enabled Businesses: A Study of the Online Brokerage Industry. *Information Systems Research*, 255-274.
- Pfeifer, P. E. (2005). The optimal ratio of acquisition and retention costs. *Journal of Targeting, Measurement and Analysis for Marketing*, 13(2), 179-188.
- Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, 81(12), 46-55.
- Reichheld, F. F., & Covey, S. R. (2006). The ultimate question: Driving good profits and true growth (Vol. 211). Boston: *Harvard Business School Press*.
- Schneider, D., Berent, M., Thomas, R., & Krosnick, J. (2008, June). Measuring customer satisfaction and loyalty: Improving the 'Net-Promoter'score. *Poster presented at the Annual Meeting of the American Association for Public Opinion Research, New Orleans, Louisiana*.

Shah, D., Rust, R. T., Parasuraman, A., Staelin, R., & Day, G. S. (2006). The path to customer centricity. *Journal of service research*, 9(2), 113-124.

Zaki, M., Kandeil, D., Neely, A., & McColl-Kennedy, J. R. (2016). The fallacy of the net promoter score: Customer loyalty predictive model. *Cambridge Service Alliance*, 10, 1-25.