



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

SEGMENTACIÓN DE CLIENTES SEGÚN SU RECEPTIVIDAD A CAMPAÑAS DE
MARKETING DE PRODUCTOS FINANCIEROS MEDIANTE EXPERIMENTOS
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

MATEO COVACICH PALET

PROFESOR GUÍA:
MARCEL GOIC FIGUEROA

PROFESOR CO-GUÍA:
DANIEL SCHWARTZ PERLROTH

COMISIÓN:
CAROLINA SEGOVIA RIQUELME

SANTIAGO DE CHILE

2020

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE: INGENIERO CIVIL INDUSTRIAL
POR: Mateo Covacich Palet
FECHA: 28/09/2020
PROFESOR GUÍA: Marcel Goic Figueroa

RESUMEN EJECUTIVO

El trabajo se sitúa en una empresa de retail financiero, en la Subgerencia de Campañas y Data Quality, que utiliza herramientas para definir los clientes que reciben campañas publicitarias de distintas índoles relacionadas con la empresa, entre otras tareas.

El problema identificado en la operación del área fue que no se conocía la efectividad de las campañas que esta subgerencia asigna. Dado esto, se pierden recursos en campañas poco efectivas. Se propone entonces un análisis de efectividad de campañas que permita hacer un perfilamiento de todos los clientes, usando como criterio cuán influenciables son por las campañas actuales de marketing de la empresa.

La metodología es Uplift Modelling, que cumple con el objetivo de analizar el efecto de campañas de algún tipo en clientes, ya que se basa en medir la diferencia en efecto entre los clientes que fueron expuestos a tales campañas y los que no, estimando un indicador que permite diferenciarlos.

Con el fin de obtener los datos para lograr crear el modelo de Uplift y categorizar a los clientes según esta práctica, se lleva a cabo un experimento de campo. Consiste en seleccionar una muestra aleatoria de clientes que cumplan ciertas condiciones, y dirigirles campañas de marketing para evaluar los resultados. Con los resultados de este experimento se procede a crear, entrenar, validar y utilizar un modelo de estimación del indicador en otros clientes. En el análisis de resultados se concluye que es mejor resultado publicitar a la mayoría de los clientes, pero no siempre a través de todos los canales disponibles.

Luego, se propone un nuevo sistema de targeting de los clientes para las campañas empleadas, usando los resultados del modelo. Se procede a hacer una evaluación de la aplicación del sistema usando los datos disponibles de campañas reales aplicadas en meses posteriores al experimento, obteniendo resultados positivos comparados con una asignación aleatoria y con el sistema actual de la empresa. Además, se descubre una relación positiva de la receptividad con la rentabilidad y comportamiento de pago de los clientes, que impulsa el uso de este sistema de priorización.

AGRADECIMIENTOS

El proceso del proyecto de memoria fue largo y accidentado. Un año en el que las condiciones cambiaron en maneras en las que jamás habría podido anticipar al iniciar en agosto del 2019. Sin embargo, ha sido un viaje que me ha enseñado cosas sobre mí y sobre el mundo que no pudiera haber aprendido de otra manera.

Al llegar al final de esta carrera no puedo dejar de pensar en todas las personas maravillosas que han sido una influencia en mi crecimiento y aprendizaje durante los años, personas para las que dedico unas palabras en este párrafo, que se quedan muy cortas para agradecer su real impacto en mi vida.

Gracias a mi familia, madre, padre y hermana, por ser el sostén tan necesario de una vida universitaria agitada y muy difícil al inicio, por preocuparse por mí y ayudarme cuando no sabía que lo necesitaba.

Gracias a mis compañeros de equipo de vóleibol de las selecciones de la facultad y de la universidad, Matías, Bastián, Renato, Vicente, Javier, Ignacio, Héctor, Diego, Sven, Pablo, Felipe, Joaquín, Ricardo, Rodrigo, Gabriel, Franco, y todos los demás, que junto con ellos y en las canchas aprendí tantas cosas acerca de lo que significa ser deportista y ser parte de un equipo de personas. Gracias también a los entrenadores que tuve, Jorge, José, Juan, Miguel y los otros que también participaron en mi carrera de deportista universitario.

Gracias también a Manuel, por enseñarme a dejar inseguridades atrás y a ser uno mismo. A Catalina, por enseñarme que la amistad significa mucho más que el contacto constante. Gracias también a Carla y Mathias por compartir y apoyarnos compartiendo vida universitaria y aún más en este último tramo.

Gracias a Javiera, con quien compartí la mayor cantidad de tiempo en estos años, por ser constante apoyo, crecimiento, cariño, y por habernos convertido juntos en las personas que somos ahora, aprendiendo muchísimo de nosotros mismos en el camino.

Gracias a todos mis compañeros del equipo docente del IN3001, que fue el espacio en el que realmente aprendí más que nunca, donde me tiré a la piscina y aprendí no solo a crecer como persona sino que también a aportar lo que pueda lograr que otros crezcan. Gracias a Camila, Romina, Kay, Montserrat, Josefa, Carla, Tomás, Cristóbal, Franco, Carlos y todos los demás que estuvieron en mi paso por ese tremendo espacio.

Finalmente, quiero dar las gracias a todos los profesores y miembros de equipos docentes que daban todo de sí para que nosotros como estudiantes aprendiéramos por sobre todo, dejando aparte otras cosas y poniendo el esfuerzo en nosotros con una preocupación y dedicación que se nota.

TABLA DE CONTENIDO

RESUMEN EJECUTIVO.....	II
AGRADECIMIENTOS.....	III
TABLA DE CONTENIDOS.....	IV
ÍNDICE DE FIGURAS	V
ÍNDICE DE ECUACIONES.....	V
ÍNDICE DE TABLAS	V
ÍNDICE DE GRÁFICOS	VI
INTRODUCCIÓN.....	1
CUERPO	2
1. CARACTERIZACIÓN DE LA EMPRESA	2
2. DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN	4
2.1 Información del área.....	4
2.2 Definición de targets de campañas	4
2.3 El problema y la oportunidad	7
2.4 Alternativa de solución	9
2.5 Objetivos.....	11
3. ALCANCES	13
4. MARCO CONCEPTUAL.....	15
4.1 Uplift modelling	15
4.2 Medida de receptividad o LIFT	16
4.3 Experimento de campo.....	17
4.4 Conditional Average Treatment Effect (CATE).....	18
4.5 Estimación de LIFT	19
5. DISEÑO EXPERIMENTAL	21
5.1 Condiciones experimentales.....	21
5.2 Selección de muestra.....	23
5.3 Fase experimental	25
5.4 Análisis de datos	25
5.5 Comprobación de hipótesis y conclusiones.....	26
6. ANÁLISIS EXPLORATORIO	27
7. RESULTADOS EXPERIMENTALES.....	33
8. MODELO UPLIFT	36
8.1 Construcción.....	36
8.2 Evaluación de modelo	37
8.3 Construcción de modelo específico para canales	43
8.4 Aplicación de modelo a campaña real.....	47
9. ANÁLISIS DE RESULTADOS.....	52
9.1 Comportamiento de pago de clientes	52

9.2 Rentabilidad de clientes.....	56
9.3 Montos y costos	57
9.4 Comportamiento de LIFT	59
 CONCLUSIONES.....	 62
 BIBLIOGRAFÍA.....	 65
 ANEXOS.....	 65
Anexo 1: Estadísticas de LIFT por decil de uplift bins.	66
Anexo 2: Resultados modelo predicción de pago de deuda	67
Anexo 3: Matriz de confusión datos de entrenamiento.....	68
Anexo 4: Matriz de confusión datos de testeo	68

ÍNDICE DE FIGURAS

Figura 1: Definición final de segmentos y canales para campañas.....	5
Figura 2: Clasificación de clientes en deciles según LIFT.	20
Figura 3: Pasos de la metodología experimental.	21

ÍNDICE DE ECUACIONES

Ecuación 1: Cálculo de LIFT y de SCORE.....	16
Ecuación 2: Cálculo de CATE.	18
Ecuación 3: Criterios de divergencia.....	20
Ecuación 4: Cálculo del éxito de la campaña.	23
Ecuación 5: Cálculo de Tasa de Respuesta Sleeping Dogs	50
Ecuación 6: Estimación del efecto sleeping dogs en la tasa de respuesta.	50

ÍNDICE DE TABLAS

Tabla 1: Variables para modelo predictivo.....	28
Tabla 2: Estadísticas porcentuales de variables demográficas de sujetos seleccionados para los experimentos de febrero y marzo.	29
Tabla 3: Distribución de tarjetas adquiridas dentro de los grupos experimentales de febrero y marzo.	31
Tabla 4: Resultados experimentales de ambos experimentos.	33
Tabla 5: Lista de variables utilizadas para entrenamiento de modelo de Uplift Random Forest.	37
Tabla 6: Importancia de las variables para el modelo entrenado.	38
Tabla 7: Tabla ejemplo de cálculo de compras incrementales para curva de Qini	41
Tabla 8: Estadísticas de LIFT por decil de uplift bins para modelo de canales.....	44

Tabla 9: Comparación de tasas de respuesta tomando distintos porcentajes de población ordenada según los tres sistemas distintos.	49
Tabla 10: Comparación de resultados de cada sistema de priorización de campañas según porcentaje de la población contactado.	50
Tabla 11: Comportamiento de pago de deuda de clientes según decil de LIFT.	53
Tabla 12: Variables de entrenamiento de modelo de predicción de probabilidad de pago. ..	53
Tabla 13: Resultados de modelo de predicción de pago de deuda por decil de LIFT.	55
Tabla 14: Análisis de rentabilidad de clientes por deciles de LIFT.	56
Tabla 15: Monto acumulado de ventas de SAE de campañas de enero según porcentaje de la población contactada, en millones de pesos.	58
Tabla 16: Costos de campañas acumulados para campañas de SAE en enero según porcentaje de la población contactada, en millones de pesos.	58
Tabla 17: Análisis de distribución de valores de LIFT en los clientes de campañas de enero y diciembre.	60
Tabla 18: Resultados de sistema de priorización por LIFT en campaña de enero, con el cálculo de la diferencia en TR al pasar de un escalón al siguiente.	61

ÍNDICE DE GRÁFICOS

Gráfico 1: Estadísticas de canales seleccionados en campañas comunes en promedio.....	14
Gráfico 2: Distribución de rango etario en sujetos experimentales de febrero y marzo.....	30
Gráfico 3: Distribución de los cupos máximos de las tarjetas de los clientes de los grupos experimentales de enero y febrero.....	31
Gráfico 4: Baldes de uplift para LIFT calculado a partir del experimento	40
Gráfico 5: Curva de Qini del modelo entrenado.....	42
Gráfico 6: Baldes de uplift para LIFT calculado con ambos modelos.....	43
Gráfico 7: Curva de Qini para el modelo de comparación de tratamientos Todo-canal v/s Sólo-call.. ..	45
Gráfico 8: Curvas de Qini para los modelos de comparación de tratamientos Todo-canal v/s Sólo-call (Gráfico 7) y Todo-canal v/s No-call	46
Gráfico 9: Baldes de uplift para LIFT calculado en clientes de campañas de enero.....	48
Gráfico 10: Curva de Qini de clientes de campañas reales de enero.. ..	48
Gráfico 11: Gráfico de resultados de sistemas de priorización según porcentaje de población contactado.....	51
Gráfico 12: Curva ROC de modelo de predicción de pago de deuda.	54

INTRODUCCIÓN

La empresa es una alianza entre uno de los mayores bancos del mundo y uno de los retailers más grandes de Sudamérica, que tiene como producto principal la tarjeta de crédito de su mismo nombre, y los productos financieros asociados a ella. La empresa tiene alrededor de 1.8 millones de clientes a lo largo de todo el país, y su público objetivo es toda persona chilena mayor a 25 años. Desde este año, se está pensando un plan piloto para incorporar a los extranjeros residentes en Chile, e incluso un programa para niveles internacionales respecto a la tarjeta de crédito.

En 2019, la empresa tuvo cerca de 4 millones de transacciones, evaluadas aproximadamente en 133 mil millones de pesos, alcanzando un 8,33% del total del mercado financiero. Este mercado está liderado por Banco Santander-Chile (32,96%), Banco de Chile (20,91%) y Banco del Crédito e Inversiones (10%). Los competidores más cercanos a la empresa son Banco Ripley (8,64%) y Banco Itaú-Corpbanca (5,13%). En un mercado relativamente estable, ha sido la empresa del rubro con mayor crecimiento anual respecto a la competencia y tienen actualmente el segundo programa de fidelización más grande de Chile.

En el último tiempo la empresa ha dado un giro en su estrategia hacia una mentalidad de foco en el cliente y transformación digital. En relación con este cambio, una actividad clave para la empresa son sus campañas de marketing dirigido a sus clientes. El proyecto de memoria presentado en este informe tiene como objetivo aportar a alinearse con la estrategia mediante las campañas de marketing de la empresa.

Para lograr poner el foco en el cliente y adoptar el mejor enfoque para cada uno de ellos, es importante conocerles, recopilar sus datos y saber utilizarlos para concluir y poder ejecutar las mejores acciones.

Una duda importante con relación directa a las campañas de marketing es si a los clientes les afectan las campañas. Saber a qué clientes impactan las campañas y si ese impacto es positivo o negativo parece crucial al momento de escoger el sistema de targeting. A pesar de esto, al tener la empresa casi 2 millones de clientes y múltiples campañas que se implementan al mismo tiempo a través de múltiples canales, se hace complicado adquirir ese conocimiento. En este proyecto, usando los datos disponibles para cada cliente, se busca estimar esta receptividad y su relación con el canal a través del que se hace llegar la campaña a través de una metodología de Uplift alimentada por experimentos de campo.

CUERPO

1. CARACTERIZACIÓN DE LA EMPRESA

El principal producto de la empresa es la tarjeta de crédito del mismo nombre. La tarjeta de crédito se define como un contrato entre la empresa y el cliente, materializado en un documento de material plástico para efectuar compras sin pagos en efectivo, e incluso incurrir en pagos futuros respecto al mismo bien. Suelen tener un límite de dinero para gastar en productos y servicios (el cupo), y la entidad emisora de la tarjeta puede cargar al cliente un porcentaje por el servicio (el interés) y, en algunos casos, una cuota final anual.

Las tarjetas de crédito suelen tener un límite de dinero que permite que la persona compre o consuma servicios hasta ese límite. Alcanzado éste, la tarjeta se inhabilita. No obstante, la entidad emisora de la tarjeta de crédito carga al comerciante un porcentaje por este servicio y en algunos casos una cuota fija anual al tenedor.

Los créditos de consumo son préstamos que las instituciones financieras dan para que el cliente pueda hacerse con el bien que necesite. Suelen emplearse también si el poseedor necesita dinero en efectivo disponible al corto plazo. En general tienen mayor tasa de interés debido a la inmediatez con la que son requeridos.

Los avances en efectivo, uno de los distintos productos financieros, son las operaciones de la tarjeta de crédito que te permiten obtener dinero en efectivo, ya sea a través de un traspaso de límite de crédito o desde la misma cuenta. Al igual que los créditos de consumo, poseen una tasa de interés alta debido a la instantaneidad con la que se pide el dinero, por lo que se recomienda elegir el pago en pocas cuotas para no aumentar la deuda.

Como último producto se tienen los súper avances, otro producto financiero. Estos consisten en un sobregiro extra en la tarjeta de crédito para pedir efectivo de manera adicional, todo esto más allá del límite que se establezca entre ambas partes. El súper avance puede tener un número variable de cuotas mensuales, acordado a por el cliente y la empresa, siendo un número común el de 12 cuotas, correspondientes a un año. Los intereses serán mucho mayores respecto al cupo inicial.

La empresa atiende a sus clientes presencialmente en sus "espacios financieros" que funcionan como sucursales y están localizados dentro de espacios de retail, como puede ser una tienda en un centro comercial. Estos

espacios financieros están presentes en muchas tiendas asociadas a la empresa a lo largo de todo el país.

La ventaja competitiva que la empresa declara tener es desde la perspectiva de la información de sus clientes. Se puede sacar mucho provecho del hecho de tener los datos de sus clientes tanto del sistema financiero como de las transacciones de la tarjeta y de las tiendas de la empresa, cosa positiva para la alianza.

2. DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN

2.1 Información del área

El área de la empresa en donde se desarrollará el proyecto es la Subgerencia de Campañas y Data Quality, parte de la Gerencia de Inteligencia de Clientes (conocida también por el nombre de Business Intelligence o BI). Esta área nació por la necesidad de dar una experiencia de compra satisfactoria al cliente y tener un registro del proceso de su compra, con el fin de obtener información sobre él.

Las principales funciones del área de Campañas y Data Quality son:

- Analizar los perfiles de los clientes para definir los clientes target de cada campaña.
- Optimizar los datos de contacto de los clientes a través de las campañas.
- Definir el segmento y los canales por los que se transmitirán las campañas publicitarias.
- Evaluar la efectividad de sus campañas.

Las decisiones que toma la gerencia de Inteligencia de Clientes son siempre respecto al comportamiento pasado de los clientes y a las acciones que se tomaban en el pasado con las campañas. Es decir, todos los resultados que se le entregan a la gerencia de Marketing y Segmentos van desde la base de que no se sabe con exactitud lo que está gatillando el éxito de una campaña sobre otra.

Dado esto, no se está aprovechando al máximo la variedad y la cantidad de los datos que se generan a través de las campañas para afinar su conocimiento de los clientes con el fin de diseñar los leads de las campañas de la mejor manera. El proyecto se hace cargo de esta parte del problema identificando preliminarmente a los clientes más propensos a responder a estímulos de marketing.

2.2 Definición de targets de campañas

Como se mencionó anteriormente, la Subgerencia de Campañas y Data Quality tiene la tarea de definir mensualmente a qué clientes les llegarán las campañas de marketing creadas por otras áreas de la empresa, y a través de qué canales

lo harán. Esta tarea supone la mayor parte del tiempo de trabajo del área, y es definida como la más importante, al tener que manejar las campañas de todos los productos y tipos.

Actualmente, para definir qué clientes se asignan para recibir publicidad de cada tipo cada mes y a través de qué canales, se utilizan distintos criterios, que se detallan a continuación, en una descripción del proceso AS-IS.

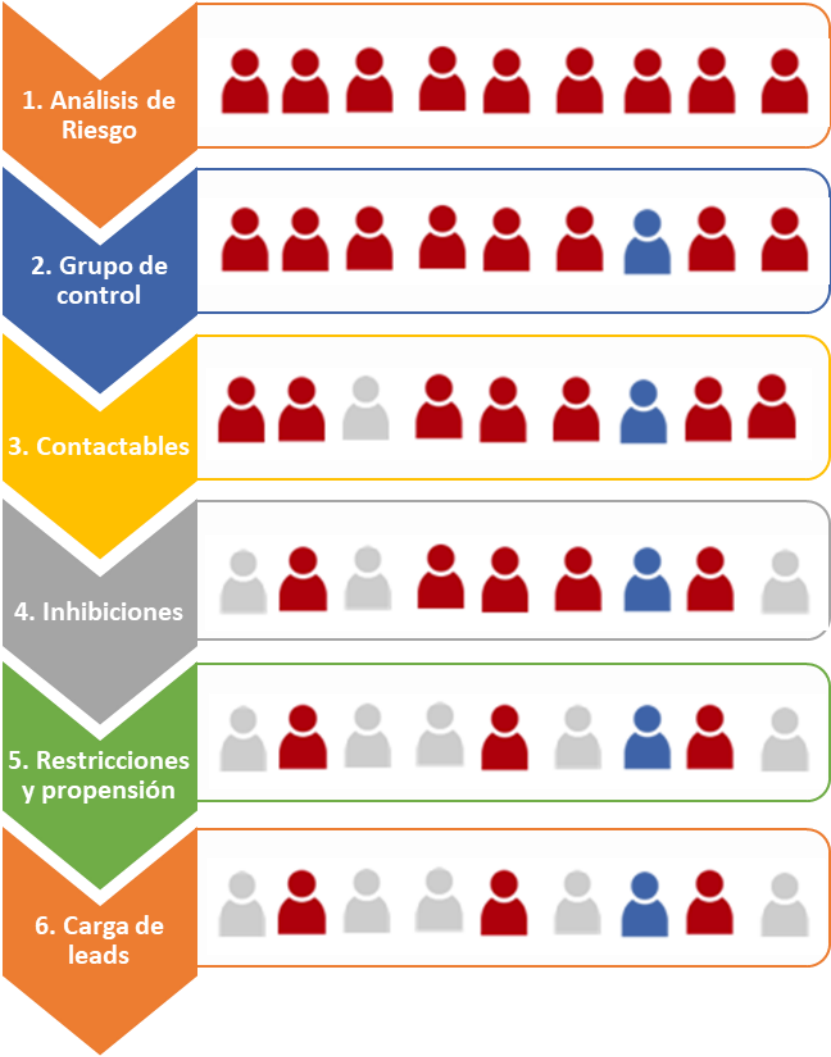


Figura 1: Definición final de segmentos y canales para campañas.

1. **Riesgo:** De toda la base de datos de clientes, el área de Riesgo de la empresa envía una lista todos los meses con los clientes que están aprobados para poder contratar ciertos productos, por ejemplo, el Súper Avance en Efectivo. Además, envían una predicción de propensión del cliente a contratar tal producto. Este es el primer filtro, pues el cliente

que no esté en la lista aprobada de Riesgo para el producto X no será contactado con una campaña de tal producto.

2. Grupo de control: Desde agosto del 2019, como política del área, se estableció como estrategia definir un grupo de control distinto cada mes para cada tipo de campaña, de un número determinado de clientes, con el fin de poder compararlo con los clientes contactados y lograr aprendizajes de esta comparación. Este grupo no recibe ninguna comunicación de su grupo determinado, pero sí puede recibir de otros. Por ejemplo, si un cliente está en el grupo de control del Avance en Efectivo, no recibirá ninguna campaña de ese producto ese mes, pero puede recibir campañas de seguro automotriz u otras. Este grupo de control se elige aleatoriamente cada mes entre los clientes que están en la lista de riesgo para cada tipo de campaña.
3. Contactabilidad: Luego, se analiza la contactabilidad del cliente por canal, es decir, si se tiene su número de teléfono, dirección de e-mail o dirección domiciliaria. Esta es otra restricción para elegir clientes para los canales, ya que no se puede enviar SMS o llamar a quien no se tenga su teléfono, y la empresa no tiene los datos completos de todos sus clientes.
4. Inhibiciones: El siguiente filtro aplicado contiene a todos los clientes que han presentado reclamos al SERNAC o a Servicio al Cliente de la empresa alguna vez acerca de las comunicaciones de la empresa. Estos clientes son removidos de la lista de posibles clientes de contacto.
5. Restricciones: Para algunos productos, como los créditos, existe una barrera de monto o cupo mínimo. Si el monto mínimo que calcula Riesgo para el crédito de un cliente está bajo un cierto umbral, ese cliente no podrá ser contactado. Esta restricción es por canal, para el call-center el monto debe ser sobre \$500.000 y para SMS y e-mail debe ser sobre \$300.000.
6. Propensión: En Riesgo, se calcula para cada cliente un score de propensión, que representa la probabilidad del cliente de contratar un producto dado. La forma de calcular este modelo se maneja desde el área de Riesgo, y no es información compartida con Campañas y Data Quality, por lo que no se conoce el modelo. Aún así, se sabe que se entrega un número que es el score, y que según ese número el cliente es categorizado como verde, amarillo, rojo o morado, donde verde es el mejor scoring y morado el peor. Esta medida por colores es actualmente la manera que tiene el área para priorizar a sus clientes para asignarlos a sus campañas. Priorizan primero los clientes con mejor medida de propensión.

Luego de este proceso de filtrado del universo de clientes disponibles, se procede a la carga de leads, es decir, a elegir específicamente qué clientes recibirán qué campañas a través de cada canal y asignarlos en las bases de datos correspondientes para que sean aplicadas. Actualmente la influencia más grande al priorizar es el scoring definido por Riesgo, siendo esta la métrica que está definiendo la asignación.

2.3 El problema y la oportunidad

En el área de Campañas y Data Quality surge la necesidad de hacer distinciones entre sus clientes en cuanto a las campañas publicitarias que se les dirigen. Antes del cambio de foco se usaba una campaña de marketing en la que se enviaban las campañas a todos los clientes seleccionados usando todos los canales disponibles indiscriminadamente. En el momento actual se diagnosticó esa estrategia como obsoleta, y se decide cambiar a un enfoque de usar el canal de marketing más adecuado y efectivo para cada tipo de cliente. Sin embargo, surge el problema de que la empresa no tiene seguridad con respecto a la efectividad de sus campañas o de la secuencialidad de sus canales.

Esto lleva a que no puedan definir las campañas óptimas para cada tipo de cliente, y que no se estén aprovechando al máximo los recursos de la empresa, tanto en costos de los canales en sí como en costo de horas-hombre de los trabajadores del área, ya que se requiere mucho trabajo para definir los envíos de las campañas, y al no hacerlo de manera efectiva se generan esfuerzos perdidos.

Preliminarmente se creía que este era el problema, entonces nace la oportunidad de hacer un estudio de los canales de marketing disponibles y determinar su efectividad a través de una serie de experimentos de campo usando las campañas de la empresa en los segmentos de clientes.

Sin embargo, tras un estudio más en detalle del funcionamiento de las asignaciones de los clientes a los canales de marketing, se descubre que el problema de los canales es en realidad un síntoma de un problema más importante, y que es el que causa la incertidumbre real: el no saber el efecto que tiene el hecho de estimular en sí, independientemente del canal, en los clientes, y las diferencias que hay entre unos y otros.

Más importante que definir qué canal es el más efectivo para estimular a un cliente, es entender si el esfuerzo de marketing, el estímulo en sí, es efectivo para el cliente. Es decir, si el cliente efectivamente se ve impactado

positivamente por el hecho de recibir la campaña o no, y qué tipos de clientes son los que reaccionan.

Esta oportunidad va aún más alineada con la estrategia mencionada anteriormente, dirigida a conocer al cliente, y saber si es valioso para el cliente y para la empresa hacerle llegar campañas de marketing. Como agregado, teniendo este conocimiento, se pueden enfocar las campañas de marketing únicamente en los clientes en que se demuestra que sí tienen impacto, y ahorrarse todos los costos de campaña en los clientes en los que no.

De manera resumida, se presentan los antecedentes del tema de memoria:

- **Necesidad de alinear con la visión estratégica:** La Subgerencia se ve en la necesidad de dirigir la asignación de campañas con foco en el cliente, y para eso necesitan conocer el efecto de éstas en cada cliente.
- **Reducción de presupuesto:** En la subgerencia se ha reducido el presupuesto para enviar campañas, por lo que se debe hacer el trabajo más eficiente. El hecho de seguir haciendo campañas caras que pueden no ser efectivas impacta directamente en este punto
- **Decisiones en base a datos:** Se espera del área de Inteligencia de Clientes que logre conocer a cabalidad el comportamiento de los clientes, y que tomen sus decisiones basadas en los datos que manejan de ellos, es decir, que funcionen como IDO (Insight Driven Organization). IDO significa ser una organización impulsada por el conocimiento relacionado con el análisis de datos y el razonamiento de la toma de decisiones, todos los días. Así, debe descubrir cómo escalar en los proyectos de toda la organización para generar un mayor impacto comercial.^[1]

Las potenciales -y actuales- consecuencias de esta oportunidad desperdiciada por parte de la empresa son las siguientes:

- **Realización de campañas dirigidas a clientes:** Con el supuesto dicho anteriormente, no se sabe si se está llegando al público que mejor responderá a las campañas de marketing, ni de la manera más efectiva. Dicho esto, hay que sacar mejor provecho de lo que se está haciendo y con esto impulsar a las campañas hacia mejores resultados.
- **Pérdida de potenciales clientes:** Al no comunicarse con el cliente a través de los mejores medios, o al no estar comunicándose con los clientes correctos, se están perdiendo oportunidades de que hagan uso de las ofertas y campañas de la tarjeta de crédito, y con esto se pierden ingresos importantes.

- **Visión de la empresa:** El proceso de transformación digital de la empresa es global. Se espera que cada parte y área de la empresa se contextualice bajo los mismos términos, por lo que este enfoque se alinearía con poner el foco centrado en el cliente y en sus necesidades e intereses desde Campañas y Data Quality sería un gran paso hacia el alineamiento con la visión de la empresa.

Tomando en cuenta todos estos puntos de impacto amplio, se propone una alternativa que busca generar una herramienta para generar efectos positivos en la búsqueda de estos puntos.

2.4 Alternativa de solución

Al estar dirigiendo las campañas por asignación según su medida de propensión, y tomando como supuesto que esta métrica está bien definida y probada, se estarían dirigiendo las campañas a los clientes con mejor scoring, es decir, con la mejor probabilidad de tomar el producto en cuestión. Sin embargo, la empresa no sabe si estos clientes son los mismos clientes que tienen el mejor efecto marginal positivo de la intervención, es decir, de la campaña en sí.

En su trabajo *Retention Futility: Targeting High-Risk Customers Might Be Ineffective*, Ascarza (2018) estudia la estrategia de retención de clientes basada en llevar a cabo una intervención activa a los clientes identificados como los más propensos a abandonar el producto o servicio que consumen (a fugarse). Ella postula que los clientes con una propensión de fuga muy alta probablemente van a hacerlo de todas maneras aunque se intervenga, por lo que sería más efectivo detectar qué clientes tienen la mayor probabilidad de responder positivamente al tratamiento y concentrar los esfuerzos en estos clientes. Luego, procede a demostrar que comúnmente los clientes con mayor propensión no necesariamente son los mismos que tienen la mayor probabilidad de responder positivamente a la intervención, por lo que se pierden recursos de la empresa en hacer llegar campañas de marketing a los clientes con mayor propensión a abandonar y que tienen pocas probabilidades de responder al tratamiento.^[2]

Desde el estudio de Ascarza se puede hacer la aproximación al caso de hacerles llegar un tratamiento en forma de campañas de marketing a clientes con el fin de que contraten un servicio o adquieran un producto. Suponiendo un caso en el que el valor del producto es el mismo en todos los casos, lo óptimo al lanzar estas campañas de marketing, dado que implementarlas supone un costo para la empresa, sería dirigir las priorizando a los clientes con

la mayor probabilidad de responder positivamente a estos estímulos de marketing. Dicho de otra manera, si se diseña una campaña para el 30% de todos los clientes de la base de la empresa, se debería elegir al 30% con mejor probabilidad de responder positivamente al tratamiento.

Como se mencionó previamente, es común que estos clientes más receptivos no sean los clientes que tienen la mayor propensión, por lo que se hace necesario identificarlos de otra manera. En el caso de lograr identificar qué clientes tienen estas mejores probabilidades de responder positivamente al marketing, se podría cambiar el sistema actual de priorización de clientes para las campañas usando este indicador. De esta manera se busca optimizar el impacto de las campañas en los clientes.

Por otra parte, es conocido por el área que, entre sus distintos canales para hacer llegar sus campañas a los clientes, el canal que tiene un mayor índice de éxito es el de call-center. Sin embargo, a un cliente le pueden llegar comunicaciones de una campaña a través de más de un canal y en más de una ocasión. Una pregunta que es interesante de contestar es si, una vez definido que un cliente será contactado por llamada telefónica, hacerle llegar publicidad por otros canales aumenta sus probabilidades de contratar el producto o no.

Esta pregunta presenta otra oportunidad que puede ser abordada en el proyecto de memoria a través de los mismos métodos que se plantean para el foco ya descrito. La hipótesis inicial que se plantea ante esta incógnita es que efectivamente las comunicaciones a través de múltiples canales tienen un efecto positivo en las campañas distribuidas por call-center. Luego, comprobada esta hipótesis, se busca analizar si vale la pena desde un punto de vista costo-beneficio agregar publicidad multi-canal a los clientes que van a ser tratados a través de un call-center.

Para lograr estimar las probabilidades de respuesta positiva a las campañas y el impacto de los demás canales se proponen 3 etapas:

1. Un experimento de campo piloto para recabar datos de respuestas de clientes a una campaña de marketing determinada, usando una muestra suficiente de clientes, y separándola en un grupo de tratamiento que recibiría las campañas, y un grupo de control que no. Adicionalmente, el grupo de tratamiento se divide en dos tipos de tratamiento con distintos canales.
2. El entrenamiento de un modelo predictivo que utilice los datos del piloto para calcular la probabilidad de responder a las campañas para todos los clientes.

3. Evaluar el posible impacto de utilizar un sistema de priorización basado en el modelo usando los datos existentes de las campañas de la empresa.

En resumen, la alternativa de solución propuesta es un análisis de efectividad de campañas de marketing según cliente a través de experimentos de campo, usando las campañas usuales de publicidad de la empresa, que son manejadas por la Subgerencia de Campañas y Data Quality.

Este trabajo entregaría resultados cuantitativos a un cierto nivel de confianza respecto a la efectividad de las campañas de la empresa por cliente, y con esta información asimilada, Campañas y Data Quality podría diseñar campañas optimizadas para los clientes que son influenciados por estas, y así aumentar la probabilidad de que una campaña tenga resultados positivos, además de no gastar recursos implementando campañas en los clientes que no responden a los estímulos de marketing. Los resultados serían confiables dada la metodología experimental que se basa en la teoría de que la única variable intercambiada entre un grupo u otro es el hecho de recibir la campaña, manteniendo una selección aleatoria de los individuos.

Dado esto, se definen en la siguiente sección los objetivos del proyecto.

2.5 Objetivos

Objetivo general:

Diseño de una metodología para conocer el comportamiento de los clientes de la empresa frente a intervenciones de marketing a través de distintos canales, con el fin de obtener una medida de receptividad de campañas individual.

Objetivos específicos:

- i. Implementar una campaña experimental piloto para los grupos definidos, y de esta manera, poder obtener datos para el modelo.
- ii. Obtener una medida de receptividad para todos los clientes a través de un modelo predictivo que permita categorizar a los clientes según receptividad a las campañas.
- iii. Comprobar o rechazar la hipótesis de que hacer publicidad por más canales además de sólo call-center tiene un impacto positivo en los resultados de las campañas.
- iv. Validar la medida encontrada a través de la estimación en clientes distintos y un análisis posterior usando datos de campañas reales de la empresa.

- v. Dejar como recomendación un nuevo sistema de priorización de targeting de clientes según una segmentación de receptividad (LIFT).

3. ALCANCES

En el proyecto de memoria se acotará el alcance del análisis de campañas de marketing a las campañas relacionadas solamente con el producto del Súper Avance en Efectivo, SAE. Este producto es el más rentable para la empresa y el segundo más contratado, por lo que su análisis es importante para Campañas y Data Quality.

Esa acotación del alcance se define debido a la cantidad de tiempo y trabajo que requiere diseñar, implementar, recopilar resultados y concluir para cada experimento, que es elevada, y se cree que acotando a este tipo de campaña el trabajo sea realizable a cabalidad. También la posibilidad de implementar experimentos se acota dentro de los primeros 6 meses del año 2020.

Fuera del análisis quedan en esta ocasión los otros productos y campañas del área, como el Avance en Efectivo (AE), las de Seguros, Traspaso de tarjeta, Segmentos, Aumento de Cupo, Créditos de Consumo, Captación de clientes, entre otros. Sin embargo, el producto SAE es muy similar al AE, por lo que se podrían extraer conclusiones muy similares al respecto de AE con las conclusiones de esta memoria.

Además, lo que realmente mide el modelo es la receptividad de los clientes ante un estímulo de marketing, comparándolos con otros clientes similares que no reciben este estímulo, por lo que el efecto incremental logra recabar solamente ese efecto.

Por otra parte, de los 8 canales de marketing que distingue actualmente Campañas y Data Quality, se analizarán sólo 4, que corresponden a los de contacto directo y de iniciativa de la empresa (outbound). Estos son:

- i. E-mail o correo electrónico
- ii. SMS o mensaje de texto telefónico
- iii. Carta impresa por correo o marketing directo
- iv. Llamada telefónica a través de *Call-center*

En el Gráfico 1 se puede observar la comparación de estos 4 canales seleccionados en cuántos clientes es común que se contacten en una campaña normal mensual, y cuántos toques por cada cliente es común realizar por canal en una campaña normal mensual.

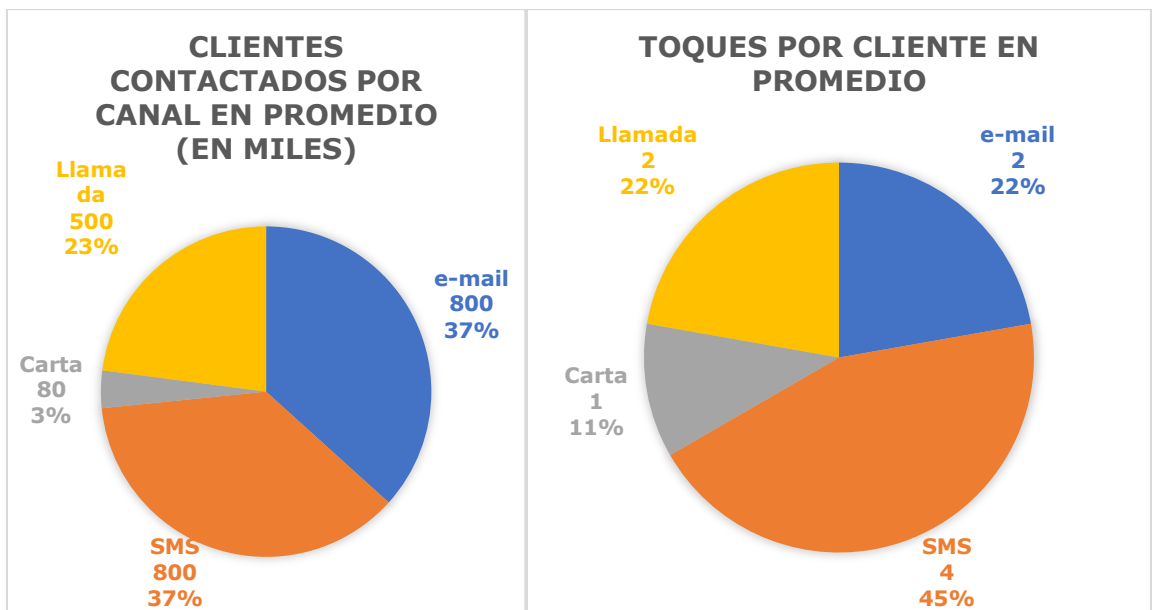


Gráfico 1: Estadísticas de canales seleccionados en campañas comunes en promedio.

Esta decisión se toma basada en los datos que se pueden obtener de esos canales que se ajustan más a las condiciones experimentales necesarias que otros. Por ejemplo, en el envío de e-mails se puede corroborar y obtener el dato de si el cliente X recibió el correo, si lo abrió y si cliqueó en algún enlace dentro de él, y eso entrega más datos para el análisis en profundidad. Similar caso es el de los otros 3 canales, si bien no se puede asegurar que el cliente efectivamente haya leído la publicidad. Distinto es el caso de otros canales, por ejemplo, el de la publicidad impresa en las boletas del retail de las tiendas de la empresa o la publicidad disponible en panfletos o anuncios en los espacios financieros, ya que no hay una manera confiable de identificar al cliente que está expuesto a este tipo de publicidad, por lo que no se puede obtener una medición confiable, por lo tanto las conclusiones que se pueden recabar no podrían alegar una causalidad directa, ya que las variables no son controlables. Mientras que en los canales escogidos se puede identificar qué cliente fue tratado con detalle individual, por lo que la medición es mucho más confiable.

De esta manera la empresa podrá sacar conclusiones para diseñar de una manera efectiva y con foco en el cliente su sistema de targeting para sus campañas de marketing, estando en conocimiento de qué clientes son más influenciables, y por lo tanto, a quienes sería mejor definir como objetivo para sus campañas de marketing.

Una dificultad que puede surgir en el desarrollo de la propuesta está en las condiciones comerciales que pueda poner la empresa a la selección de

individuos o al tratamiento de estos en cuanto a los tratamientos. Para evitar estos problemas es importante definir en detalle y con cuidado el diseño experimental.

4. MARCO CONCEPTUAL

4.1 Uplift modelling

La teoría de modelamiento en base a uplift está enfocada en la realización y evaluación de un modelo incremental, con el objetivo de medir lo que se busca en la memoria: el efecto directo y real de un estímulo de marketing.^[3] Se busca estimar este efecto a través de la medida de un indicador llamado LIFT, que es directo, es decir, a mayor LIFT, mayor es el impacto de la intervención de marketing.

Ya obtenido este indicador, la teoría sugiere segmentar la población de clientes en cuatro grupos:

1. Los *influenciables*: Estos clientes son los que son afectados positivamente por el estímulo de marketing, es decir, su probabilidad de contratar un producto aumenta si es que reciben una campaña.
2. Los *sure things* o asegurados: Estos clientes no son afectados por el estímulo, ya que se observa que independiente de haberlos estimulado o no, contratan el producto. Ya tomaron la decisión deseada, haya campaña o no.
3. Los *lost cause* o causas perdidas: Estos clientes tampoco son afectados por el estímulo, pero han tomado la decisión de no contratar el producto, independiente de si reciben la campaña o no.
4. Los *sleeping dogs* o perros dormidos: Estos clientes son los que hubieran tomado la decisión deseada de contratar, pero al recibir el estímulo de marketing deciden no hacerlo. Las campañas de marketing tienen un efecto negativo en ellos.

Esta segmentación se realiza con el objetivo de poder enfocar mejor los esfuerzos de marketing, enfocándose en los clientes clasificados como influenciables, ya que invertir recursos en cualquiera de las otras tres categorías no tendría efecto alguno, o incluso podría tener un efecto negativo no deseado por la empresa.

4.2 Medida de receptividad o LIFT

El hecho de estudiar el efecto de que un cliente responda a un estímulo de marketing en vez de estudiar el efecto de que un cliente tome un producto o no, ya ha sido estudiado en el contexto de retención de clientes (Ascarza, 2018).^[4] En este artículo se realiza un proceso similar al que se desarrollará en la memoria, usando el concepto que hasta ahora se ha descrito en este informe como la receptividad a la campaña como LIFT.

En su artículo, Ascarza define dos medidas, RISK y LIFT, y las usa para trabajar. A continuación, basándose en las definiciones del paper, se definirán las medidas SCORE y LIFT que se utilizarán en el desarrollo de la memoria.

Se considera un cliente i , con características observables X_i . Este conjunto de características es propio del cliente, por ejemplo, edad, sexo, comportamiento de compra, segmentación de la empresa, productos que posee, entre otras características demográficas o de comportamiento histórico. La práctica actual de la empresa es calcular la probabilidad de un cliente de contratar cierto producto según sus características X_i , denotado $P[Y_i | X_i]$, en donde Y_i representa el hecho de contratar el producto. Luego, se decide priorizar los esfuerzos de marketing según esta medida, priorizando las probabilidades más altas.

Sin embargo, este enfoque es limitado, ya que falla en incorporar la variable de decisión en la especificación del cálculo. La variable de decisión, denotada T_i (tratamiento), corresponde a si un cliente es elegido y recibe el estímulo de marketing (la campaña), o no.

Si el cliente i recibe la campaña, se denota con $T_i = 1$, y en caso contrario con $T_i = 0$. Esta variable T_i está completamente a decisión del área de Campañas y Data Quality, por lo tanto corresponde a la variable de decisión.

Dado esto, se definen las medidas como en la Ecuación 1.

$$LIFT_i = P(Y_i|X_i, T_i = 1) - P(Y_i|X_i, T_i = 0)$$

$$SCORE_i = P(Y_i|X_i, T_i = 0)$$

Ecuación 1: Cálculo de LIFT y de SCORE. Adaptación de ecuaciones 1 y 2 en "Retention Futility: Targeting High-Risk Customers Might Be Ineffective" (Eva Ascarza, 2018).

Se puede inferir que $SCORE_i$ es la medida actual que está calculando la empresa, la probabilidad de que un cliente tome el producto. Notar que se define que no recibe el estímulo de marketing porque la empresa calcula esta probabilidad antes de ejecutar la campaña, a diferencia del LIFT.

El argumento de Ascarza, y el que se quiere demostrar en esta memoria, es que la empresa debería enfocar sus esfuerzos de campaña en los clientes que tengan un puntaje de LIFT más alto, a diferencia de su sistema actual, en donde se enfocan en los clientes con mayor SCORE. Esto ya que es obvio que es preferible enfocarla intervención en un cliente con más efecto marginal de la misma que en uno con menor efecto pero más SCORE. De esta manera, luego se podría hacer un análisis costo-beneficio entre efecto de la intervención (LIFT) y valor del cliente.

El LIFT correspondería a la medida de receptividad buscada, es decir, la medida de un cliente de responder a un estímulo de marketing (ya sea positiva o negativamente) de parte de la empresa, ya que se compara la diferencia incremental entre los clientes que reciben el estímulo con los que no.

Si bien estimar el $SCORE_i$ es directo con los datos de la empresa, ya que no se necesita más información que la contenida en sus bases de datos, estimar el $LIFT_i$ requiere de la observación de dos escenarios que no se pueden observar al mismo tiempo, ya que no se puede tener el caso de que un cliente reciba el estímulo y al mismo tiempo no lo reciba.

Una opción para estimar el LIFT es construir dos modelos por separado, uno para un grupo expuesto al estímulo de marketing y otro para un grupo que no lo haya sido, y luego restar directamente los resultados de estos modelos. Sin embargo, este enfoque tiene la desventaja de que construye un modelo incremental a partir de dos modelos no-incrementales.

Dado esto, se decide utilizar otra opción para lograr estimar el efecto. Realizar un experimento de campo, para recopilar datos de los clientes, utilizando asignación aleatoria entre los clientes que recibirán el estímulo y los que no. Con los datos resultantes del experimento se podrá estimar el efecto incremental de la campaña, el valor del LIFT, a través de las covariables de los individuos.

4.3 Experimento de campo

El experimento de campo toma su base en definir al menos un grupo de tratamiento y un grupo de control, seleccionados de manera aleatoria de un universo definido, y aplicar un tratamiento solamente a los sujetos del grupo

adecuado, sin aplicarlo al de control. El objetivo de la selección aleatoria es eliminar los sesgos de selección y que la única diferencia significativa entre ambos grupos sea la variable del tratamiento administrado.

Se define una variable a medir, que en este caso sería el éxito de las campañas, y luego de aplicado el tratamiento y pasado el tiempo definido en el diseño del experimento, se mide esa variable y se analiza si hay suficiente evidencia entre ambos grupos para afirmar con seguridad que el tratamiento significó una diferencia o no para la variable medida.

Luego se contrasta una hipótesis nula, que generalmente indica que el efecto del tratamiento no genera una diferencia entre los dos grupos, contra una hipótesis alternativa, que indica que tal efecto existe (también puede indicar que el efecto sea mayor o menor, no solamente diferente).

4.4 Conditional Average Treatment Effect (CATE)

Siguiendo una metodología de resultados potenciales para inferencia causal,^[5] se puede asumir la existencia de los resultados potenciales $Y_i^{(1)}$ y $Y_i^{(0)}$ correspondientes respectivamente a si el cliente i habría tomado el producto con y sin el estímulo de marketing, respectivamente.

Dado esto, se puede estimar el Conditional Average Treatment Effect, o CATE, que se refiere al efecto del tratamiento, en este caso la campaña de marketing, pero considerando el set de características observables X_i de los clientes, para estimar el efecto según estas características y diferenciar entre clientes.

Luego, dado que el resultado Y_i es binario (el cliente contrata el producto o no lo hace), según la fórmula de la literatura, el CATE se expresa como se aprecia en la Ecuación 2.

$$CATE_i = \mathbb{E}(Y_i^{(1)} - Y_i^{(0)} | X_i) = P(Y_i | X_i, T_i = 1) - P(Y_i | X_i, T_i = 0) = LIFT_i$$

Ecuación 2: Cálculo de CATE.

Por lo tanto, como se puede observar, el CATE extraído de los datos del experimento entregaría el valor del LIFT buscado.

4.5 Estimación de LIFT

Para calcular el CATE, está el problema de que no se pueden observar al mismo tiempo en un individuo los resultados potenciales de que haya recibido el tratamiento y de que no lo haya recibido, por lo que comúnmente se estima ese resultado. En esta memoria se propone utilizar un modelo basado en algoritmos de Random Forest,^[6] utilizando la librería de *Uplift* en R. Para esto, el primer paso es separar todos los datos que se tengan como resultado del experimento piloto en una proporción dada, $x\%$ de datos serán categorizados como datos de entrenamiento, que serán los que le den forma al modelo, y $(1 - x\%)$ de los datos será categorizados como datos de testeo o validación, para luego comprobar la validez del modelo.

Si bien existen varios modelos propuestos para estimar LIFT en modelamientos de Uplift en la literatura, siendo común el uso de árboles de decisión,^[7] estos tienen el problema de tener una varianza muy alta, y los métodos para reducirla (como puede ser el *bagging*^[8]) están muy correlacionados entre sí, ya que usan las mismas variables predictoras.

Según Guelman, Guillén y Pérez-Marín (2015), una solución para este problema es utilizar un algoritmo basado en la técnica de Random Forest, que genera un conjunto de árboles de decisión, cada uno con una fracción v de los datos de entrenamiento del modelo. La razón para la separación de los datos es que reduce drásticamente la correlación entre los árboles del conjunto (el "bosque"). El proceso de construcción de los árboles involucra elegir aleatoriamente un número determinado n de variables del total de variables descriptivas de los datos, y usarlas como candidatas para la división de los individuos en conjuntos que luego se compararán para estimar su LIFT. El hecho de que esta selección de variables sea aleatoria disminuye aún más la posible correlación entre los conjuntos. Finalmente, el LIFT estimado final se obtiene de promediar las predicciones de LIFT de todos los árboles individuales del bosque.

Existen distintos criterios para la división de los datos dentro de los árboles de decisión que componen el bosque, y estos están basados en las llamadas distribuciones de divergencia. En esta ocasión se utilizarán las mostradas en la Ecuación 3, y se compararán sus resultados.

$$E(p_t||p_c) = \sum_y (p_t(y) - p_c(y))^2,$$

$$\chi^2(p_t||p_c) = \sum_y \frac{(p_t(y) - p_c(y))^2}{p_t(y)},$$

$$L1(p_t||p_c) = \sum_y |p_t(y) - p_c(y)|.$$

Ecuación 3: Distancia euclidiana, divergencia chi-cuadrado, divergencia L1 de Kullback-Leibler, respectivamente.

Definidos ya los criterios de división, se procede a ejecutar un algoritmo presentado en el artículo, que se puede observar en detalle en el trabajo de Guelman.^[9]

Luego de construido el modelo con los datos de entrenamiento, se procede a testearlo con el porcentaje restante de datos que no se utilizaron para su entrenamiento. Esta práctica se realiza para validar la capacidad predictiva de este y evitar el sobreajuste a los datos. De esta manera, y ya validado, está listo para convertirse en el modelo predictivo con el objetivo de predecir el LIFT de toda la base de clientes de la empresa.

Teniendo este dato, se procede a ordenar a los clientes según deciles, tomando como variable el valor del LIFT para dividirlos en un ranking de mayor a menor. Con esta definición de deciles se tiene finalmente 10 grupos de clientes ordenados de mejor a peor según su probabilidad de responder positivamente a un estímulo de marketing, como se puede ver en la Figura 2.

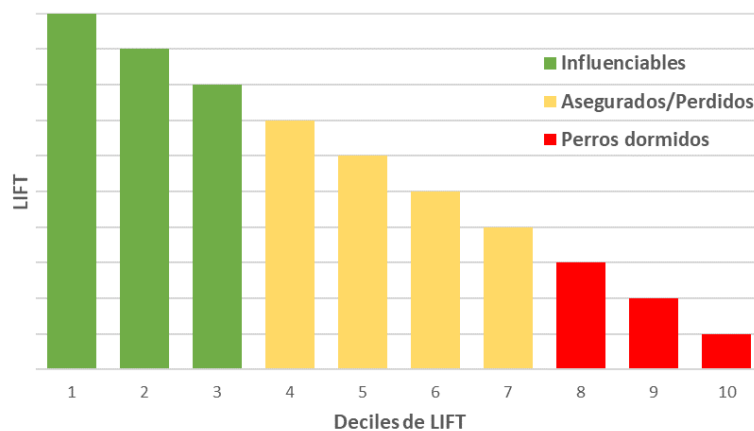


Figura 2: Clasificación de clientes en deciles según LIFT.

5. DISEÑO EXPERIMENTAL

Para el diseño experimental, se aplica la metodología experimental estadística común en experimentos de campo, que sigue los seis pasos detallados a continuación en esta sección.



Figura 3: Pasos de la metodología experimental.

Cabe mencionar que la fase de diseño de hipótesis ya se cubre en las secciones 2.3 y 2.4 de este informe.

5.1 Condiciones experimentales

En el experimento de campo se busca recopilar datos para poder estimar el LIFT. Para este experimento el diseño es simple. Consiste en elegir una muestra aleatoria de tamaño suficiente de la base de datos de clientes aprobados para un SAE (el universo definido), luego dividirla aleatoriamente en grupo de tratamiento y grupo de control, y hacer llegar una campaña publicitaria de SAE solamente al grupo de tratamiento. El grupo de control no recibe ningún estímulo del producto, y luego se comparan las contrataciones de SAE entre ambos grupos.

Dentro del grupo de tratamiento, a su vez se realiza una división aleatoria en grupo de tratamiento call-center y grupo de tratamiento Todo-canal. Los sujetos del primer grupo recibirán la campaña experimental sólo a través de llamada telefónica y los del segundo grupo a través de todos los canales

definidos en la sección 3. Luego se puede comparar los resultados entre ambos y con el grupo de control.

Como pregunta a responder, se busca ratificar que un grupo elegido en base a los mejores puntajes LIFT tendría un mejor desempeño que el sistema actual de la empresa. Sin embargo, puede que estos clientes no sean los de mayor valor para la empresa. Para abordar esta problemática, se propone calcular un indicador que se define como *Value-Uplift*, que pondera el valor del LIFT del cliente con una variable que define si es un cliente valioso para la empresa. En este caso se utiliza como indicador el hecho de si el cliente paga la primera cuota del SAE en la fecha correspondiente. Dicha variable tomaría el valor de 1 si el cliente lo hace y de 0 si cae en morosidad. De esta manera se calcularía el *Value-Uplift* como el promedio del LIFT (que toma valores entre 1 y -1) con esta variable, logrando un segundo indicador que puede ser utilizado para clasificar a los clientes y tomar decisiones de targeting.

De esta manera, además de poder clasificar a los clientes según uplift puro, también se podrían clasificar a través de *Value-Uplift*, teniendo la empresa ambos datos disponibles.

Se define como universo de sujetos a todos los clientes actuales de la empresa que hayan pasado los filtros detallados a continuación.

1. Riesgo: De toda la base de datos de clientes, el área de Riesgo de la empresa envía una lista todos los meses con los clientes que están aprobados para poder contratar ciertos productos, por ejemplo, créditos como el Súper Avance en Efectivo. Este es el primer filtro.
2. One: Se apartan 110 mil clientes de la base para experimento de One, que es otro experimento actual de la empresa sin relación a la memoria.
3. ABC1 y colaboradores: Se quitan del universo los clientes que caigan dentro de estas dos categorías, ya que el call-center se centra más en estos ya que le entregan mayor beneficio, dándoles una mayor intensidad de tratamiento con respecto a los otros. De ocurrir esto, sería perjudicial para los datos, ya que el tratamiento no sería uniforme, por lo que se opta por eliminarlos del universo.
4. Contactabilidad e inhibiciones: Se quitan del universo quienes no tengan los datos para ser contactables, y los clientes que hayan hecho reclamos al SERNAC o a Servicio al Cliente por campañas anteriores.
5. Monto mínimo: Se quitan del universo clientes con un monto mínimo de crédito menor de \$500.000, que es lo mínimo que acepta el call-center.

El resultado de estos filtros corresponde al universo del experimento. Cabe mencionar que tanto grupo de control como grupo expuesto son seleccionados

a partir del resultado de estos filtros, es decir, primero se definen los individuos del universo y luego se realiza la selección aleatoria de los grupos.

Este experimento cuenta con dos iteraciones, una durante el mes de febrero y otra durante el mes de marzo. Ambas son independientes la una de la otra y utilizan individuos diferentes.

5.2 Selección de muestra

En este paso se eligen los clientes para los grupos descritos mediante selección aleatoria para evitar sesgos de selección.

Se calcula el número mínimo de individuos necesarios para cada grupo, según lo expuesto en el marco conceptual, utilizando datos históricos y la librería *pwr* de R, que automatiza estos cálculos.

Los parámetros utilizados son:

- Orden de magnitud de la diferencia: Se estima la diferencia según datos históricos de campañas anteriores (entre junio y octubre 2019) en que se ha utilizado grupo de control. La diferencia es del orden del 0,56% entre los grupos, con una desviación estándar de un 11,3%.
- Parámetros de la variable: En este caso se define la variable *TR* según la fórmula de la Ecuación 4.

$$TR = \frac{\text{Total de productos contratados}}{\text{Total de clientes del grupo}}$$

Ecuación 4: Cálculo del éxito de la campaña. TR representa la frase Tasa de Respuesta.

- Confianza: Se define un nivel de confianza del 95%, es decir un valor de $\alpha=0,05$.
- Poder estadístico: Se define un poder 80%, es decir un valor de $\beta=0,2$.

Dados estos parámetros, se calcula que el número mínimo necesario para concluir significativamente con esos niveles de confianza y de poder es de 6.392 clientes para cada grupo. Ajustando este número por una pérdida estimada del 10% debido a condiciones extra-experimentales (práctica común y recomendada) el número crecería hasta alcanzar 7.103 sujetos, lo que significaría el mínimo necesario. Sin embargo, para asegurarse contra las pérdidas posibles, para cada grupo se tomará un número superior de individuos que el número mínimo.

Luego, y tomando en cuenta restricciones comerciales impuestas por la empresa, para el experimento piloto del mes de febrero se define un número total de 40.000 individuos, 20.000 para el grupo de control y 20.000 para el de tratamiento. A su vez, el grupo de tratamiento se divide en 10.000 individuos para el grupo Todo-canal y 10.000 para el grupo call-center. En cambio, para el experimento piloto del mes de marzo se definen 15.000 clientes para cada uno de los tres grupos, un total de 45.000. Cabe recalcar que la selección es acorde a una selección y asignación aleatoria.

Una regla a aplicar es que todos los clientes del grupo de tratamiento Todo-canal serán contactados por los 4 canales, aún cuando no lo serían en condiciones normales. Esta decisión se tomó ya que en el canal de call-center, que usualmente es el que tiene mejor tasa de conversión, el tratamiento de los clientes que se les indica que deben contactar no es homogéneo. Esto ocurre porque el call-center es externo a la empresa, y cobra su pago por venta concretada, lo que desencadena que no contacten con la misma intensidad a un cliente marcado con un buen scoring (verde, alta probabilidad de contratar el producto) que a uno con bajo scoring (rojo, morado). Usualmente llaman más y se esfuerzan más en concretar la venta del primero. Por otra parte, sus tasas son más altas al concretar ventas por montos más altos, por ejemplo, de segmento ABC1, por lo que también se enfocan más en esos clientes.

Esto supone un problema para la validez del experimento, ya que uno de los supuestos para que funcione un experimento de campo es que el tratamiento sea homogéneo en el grupo adecuado, y eso no se cumpliría si en el grupo de tratamiento hubiera clientes con categorías de scoring distintos. Podrían ser contactados con una intensidad distinta, o incluso llegar a no ser contactados. Esta diferencia de tratamiento también es la razón por la que se quitaron a los clientes ABC1 del experimento, ya que el call-center también se centra más en estos clientes que en otros.

Es por esto que la medida que se tomará será entregar los datos al call-center modificados manualmente, insertando a todos los clientes del grupo de tratamiento con una categoría de scoring verde. Además, de esto, no hay más factores externos que se correlacionen con el tratamiento, ya que ni los llamantes ni los clientes saben la diferencia entre grupos, por lo que se tratan de igual manera.

La selección de los grupos se concreta el día 20 del mes anterior al experimento correspondiente, enero o febrero según corresponda, dado a que 10 días es el tiempo que necesita el área para gestionar los contactos de las campañas para el mes siguiente.

5.3 Fase experimental

Se aplica el tratamiento diseñado a los grupos seleccionados y definidos, durante el tiempo establecido, recopilando los datos de resultados.

La implementación de las campañas se lleva a cabo de la misma manera que una campaña normal y corriente del área. La única diferencia es la selección de los grupos y las modificaciones de scoring que se mencionó anteriormente.

El tratamiento consiste en hacer llegar una campaña común del mes correspondiente a todos los clientes que estén en el grupo de tratamiento, a través de los canales que correspondan a su grupo. Los canales que entran en el estudio son e-mail, SMS, call-center, y marketing directo (carta física).

La distribución de toques para el grupo Todo-canal sería la siguiente:

- Carta: Se planifica para que el 80% de las cartas esté entregadas el primer día del mes de campaña, el 20% restante se entrega durante la primera semana.
- E-mail: Dos e-mails al mes, uno en la primera semana y otro en la tercera. Enviados los días jueves.
- SMS: Un mensaje semanal durante las 4 semanas del mes. Enviados todos los martes.
- Tratamiento por call-center.

Mientras que el otro grupo, el de tratamiento Call-center, sólo recibiría la campaña a través de las llamadas telefónicas.

El tratamiento por call-center es un servicio contratado y externo a la empresa por lo que no se puede determinar el momento en que se llame a cada cliente, eso lo decide el call-center utilizado, pero sí se puede medir y obtener el dato del momento en que esta ocurre.

5.4 Análisis de datos

En esta etapa, con todos los datos recopilados de los resultados del experimento, se revisan los resultados para luego interpretarlos y concluir.

La variable a observar de los resultados de los experimentos sería en ambos casos si el cliente contrató el SAE o no. Esta variable se llama "Venta Atribuida", y es una variable binaria que indica si la venta del SAE puede atribuirse a la campaña de marketing o no.

Las condiciones para atribuir la venta varían por canal de contacto:

- Si el cliente fue cargado a la base de datos para ser contactado por call-center y contrató SAE, la venta se atribuye, independientemente de si fue contactado o no.
- Si el cliente fue cargado a la base de datos de campaña de marketing directo (carta) y contrató SAE, se atribuye la venta.
- Para el grupo de control, se atribuye la venta si esta se produce después de la fecha mínima de acción del grupo de tratamiento. Es decir, si un cliente del grupo de control contrata SAE, se atribuye la venta sólo si la fecha es después de que hayan empezado las campañas de marketing para el otro grupo.
- Si el cliente está cargado en la base de datos de campaña por SMS, se puede medir el momento en que se envía el mensaje. Se atribuye la venta al cliente si esta ocurre después de enviado el mensaje, no antes.
- Si el cliente está cargado para las campañas de e-mail, se puede obtener información de cuando fue enviado el e-mail, y si el cliente lo recibió y lo abrió. Se atribuye la venta al cliente si contrata el producto después de haber enviado el e-mail y haber comprobado que lo recibió, aunque no lo haya abierto.

5.5 Comprobación de hipótesis y conclusiones

Con el análisis final procesado, se observa si se verifican las hipótesis iniciales o no, si es significativa la evidencia y se concluye al respecto. Al comprobarse la validez de los resultados, ya estarían dispuestos los datos para ser utilizados para entrenar el modelo para estimar el LIFT de los clientes.

Por otra parte, con los resultados de los experimentos analizados, se procede a concluir acerca de la hipótesis de que las comunicaciones multi-canal tienen un efecto positivo en las campañas distribuidas por call-center.

Lo que sería necesario para comprobar esta hipótesis sería que el grupo Todo-canal haya tenido mejor tasa de respuesta que el grupo Call-center, y que esa diferencia fuera significativa estadísticamente. Para comprobar la significancia se utilizan test de student y chi-cuadrado, frecuentemente utilizados en estas circunstancias.

6. ANÁLISIS EXPLORATORIO

La cantidad de datos disponibles de los clientes de la empresa es amplia, y de dichos datos se hace una selección de variables consideradas relevantes para el proyecto. La base de datos se arma con las siguientes columnas, mostradas y explicadas en la Tabla 1.

Como se ha mencionado anteriormente, todos los datos a utilizar se obtienen de las bases de datos que maneja el área de Campañas y Data Quality y de los experimentos. Se dispone de los datos de los clientes de la empresa, además de los datos de los clientes de ambas partes de la alianza, todos manejados desde base de datos a través de Teradata usando consultas SQL. Se trabaja en el análisis de estos datos usando Teradata y R a través de RStudio.

Nombre de variable	Categoría	Descripción
ID_RUT	ID	Identificador codificado del RUT del cliente
GC	Experimental	Valor 1 si pertenece al grupo de control, valor 0 si no
EXPUESTO_CALL	Experimental	Valor 1 si pertenece al grupo de tratamiento call-center, valor 0 si no
EXPUESTO_TC	Experimental	Valor 1 si pertenece al grupo de tratamiento Todo-canal, valor 0 si no
PERIODO	Campaña	Mes del registro
EMAIL	Campaña	Cuántos e-mails de campañas recibió en ese mes
SMS	Campaña	Cuántos SMS de campañas recibió en ese mes
MKTD	Campaña	Cuántas cartas físicas de campañas recibió en ese mes
CC	Campaña	Cuántas llamadas de campañas recibió en ese mes
NRO_TARJETAS	Tarjeta	Número de tarjetas que posee el cliente
HABILITADO_COMPRA	Tarjeta	Valor 1 si el cliente está habilitado para comprar con su tarjeta, 0 si no
HABILITADO_REPA	Tarjeta	Valor 1 si el cliente está habilitado para repactación de deuda en su tarjeta, 0 si no
TARJETAS_HABILITADAS	Tarjeta	Número de tarjetas habilitadas para hacer compras que posee el cliente
CUPO_TOTAL	Tarjeta	Cupo máximo entre sus tarjetas de crédito
DEUDA_TOTAL	Tarjeta	Máximo de deuda permitido
DIAS_MORA	Tarjeta	Días que el cliente se ha encontrado en estado de mora de pago
MONTO_SIN_PREPAGO	Tarjeta	Monto preaprobado para súper avance en efectivo

CCSAE_COLOR	Segmentación	Segmentación de campañas de call-center según SCORE
ScoreCCSAE_MR	Segmentación	Puntaje SCORE en el que se basa la segmentación SAE color
FREQ	Transaccional	Número de seguros contratados que tiene el cliente
L	Transaccional	Hace cuántos meses tomó su primer seguro
R	Transaccional	Hace cuántos meses tomó su último seguro
AVANCISTA	Transaccional	Toma el valor de 1 si el cliente tomó un avance en los últimos 6 meses, 0 si no
SEGMENTO_ACT_PPF	Transaccional	Segmentación según actividad del cliente en productos financieros en los últimos 6 meses
SEGMENTO_ACT_RETAIL	Transaccional	Segmentación según actividad del cliente en retail de la empresa en los últimos 6 meses
SEGMENTO_ACT_CONVENIO	Transaccional	Segmentación según actividad del cliente en productos o servicios no asociados a la empresa en los últimos 6 meses
SEGMENTO_ACT_SEGURO	Transaccional	Segmentación según actividad del cliente en seguros en los últimos 6 meses
SEXO	Demográfica	Sexo del cliente
TIENE_HIJO	Demográfica	Valor de 1 si el cliente tiene al menos un hijo o hija, 0 si no
CASADO	Demográfica	Valor de 1 si el cliente está casado, 0 si no
TIENE_BR	Demográfica	Valor de 1 si el cliente es dueño de al menos 1 bien raíz, 0 si no
REGION	Demográfica	Región geográfica del cliente. Se denotan con números del 1 al 16 pero se tratan como categorías.
TIENE_VEHICULO	Demográfica	Valor de 1 si el cliente es dueño de al menos 1 vehículo, 0 si no
RANGO_EDAD	Demográfica	Rango etario del cliente
Dias_antigüedad	Demográfica	Días de antigüedad del cliente en la empresa
VENTA_ATRIBUIDA	Resultado	1 si cliente adquirió SAE dentro de los márgenes del experimento, 0 si no
MONTO	Resultado	Monto del SAE adquirido, 0 si no adquiere.
CUOTAS	Resultado	Cuotas del SAE adquirido, 0 si no adquiere.

Tabla 1: Variables para modelo predictivo.

La lista anterior constituye una lista preliminar de variables. La lista final de variables que se considerarán en el modelo dependerá de su capacidad predictiva en los modelos estadísticos.

Cabe aclarar que para la variable CCSAE_COLOR existen 4 categorías según el puntaje de SCORE, verde, amarillo, rojo y morado, siendo verde la mejor, con mayor puntuación SCORE, amarillo la siguiente mejor, rojo la siguiente y morado la peor. Las variables de SEGMENTO_ACT son categorías que segmentan a los clientes que han realizado compras en los últimos 6 meses con su tarjeta en la categoría indicada. Por ejemplo, SEGMENTO_ACT_PPFF define distintas categorías para un cliente si tomó un producto financiero hace menos de 6 meses, hace más de 6 meses, o si nunca ha tomado uno.

De los sujetos escogidos aleatoriamente para los experimentos se busca analizar ciertas variables para adquirir una mirada general de su contexto y similitudes o diferencias entre ellos.

De las variables demográficas se pueden obtener las estadísticas que se muestran en la Tabla 2.

Categoría	Febrero		Marzo	
	Sí	No	Sí	No
SEXO (MASCULINO)	43%	57%	42%	58%
TIENE_HIJO	69%	31%	74%	26%
TIENE_BR	44%	56%	47%	53%
TIENE_VEHÍCULO	58%	42%	60%	40%
CASADO	51%	49%	56%	44%

Tabla 2: Estadísticas porcentuales de variables demográficas de sujetos seleccionados para los experimentos de febrero y marzo. En febrero el n total es de 40.000 y en marzo de 45.000.

De la tabla se puede concluir que en las variables demográficas binarias no hay diferencias significativas marcadas. La tendencia más grande es en el hecho de tener hijos o no, en que la balanza se inclina más hacia la respuesta afirmativa. También se pueden observar similitudes entre los grupos de marzo y febrero en todas las categorías.

Con respecto a las edades de los clientes del experimento, se puede ver la distribución en el Gráfico 2.

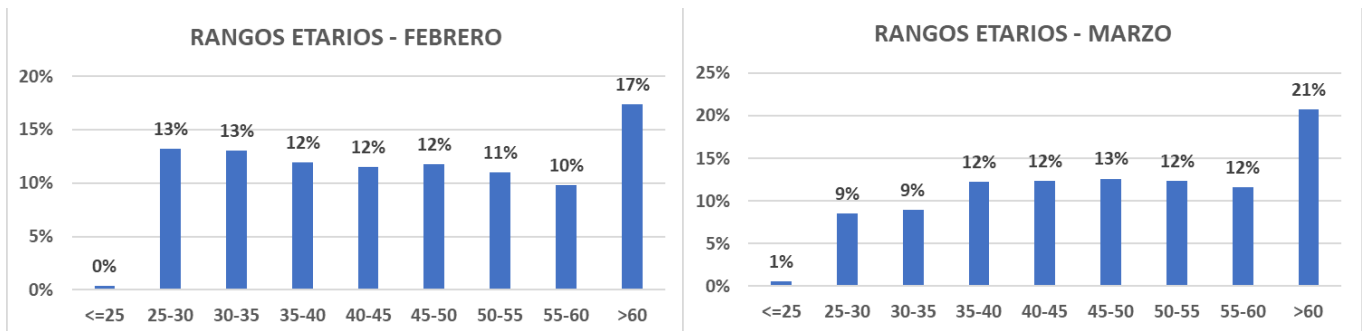


Gráfico 2: Distribución de rango etario en sujetos experimentales de febrero y marzo.

Se concluye que la distribución tiene pocas diferencias en los rangos intermedios, pero que hay un porcentaje muy pequeño de personas menores a 25 años, que se explica porque es un segmento que en proporción, tiende a tener muchas menos tarjetas de crédito. El otro extremo son las personas mayores de 60 años, que son el segmento más numeroso. Este dato se explica también en que es el rango etario en el que más clientes tiene la empresa, superando significativamente a los demás en cantidad.

Con respecto a la región de procedencia de los escogidos, en ambos grupos, febrero y marzo la región metropolitana contiene una cifra cercana al 50% de los clientes, y el resto de regiones tiene cada una menos de un 10% del número de clientes escogidos.

Con respecto a la posesión de las distintas tarjetas de la empresa por parte de los clientes escogidos para los experimentos, se puede afirmar que la gran mayoría son poseedores de una única tarjeta, un 90% en el grupo de febrero y un 94% en marzo. Los clientes que poseen dos tarjetas de la empresa son tan sólo un 5% tanto en el grupo de febrero como en el de marzo (hay aproximadamente un 5% de los datos del grupo de febrero en que aún no se ha logrado rescatar ese dato). Luego hay un porcentaje menor a un 1% de personas que tienen 3 tarjetas a su nombre, 25 personas de un total de 40.000 en el grupo de febrero, y 42 personas de un total de 45.000 en el de marzo. Como último caso, hay un cliente del grupo de marzo que posee 4 tarjetas.

Se concluye que la normalidad en los clientes de la empresa es que posean una sola tarjeta, dato que se confirma al revisar las estadísticas del resto de los clientes.

Con respecto a qué tarjetas son más y menos adquiridas, se puede apreciar esa información en la Tabla 3.

Tarjeta	Febrero		Marzo	
	Cantidad	Porcentaje	Cantidad	Porcentaje
Mastercard	30.525	77%	35.894	76%
Visa	4.115	10%	5.620	12%
Tarjeta propia	2.759	7%	3.249	7%
Mas Paris	1.102	3%	1.367	3%
Mas Jumbo	803	2%	980	2%
Mastercard Black	114	0%	176	0%
Mastercard Platinum	88	0%	165	0%
Mas Easy	65	0%	70	0%

Tabla 3: Distribución de tarjetas adquiridas dentro de los grupos experimentales de febrero y marzo.

La cantidad más grande de tarjetas en circulación las acumula la tarjeta Mastercard, seguida en menor medida por su homónima de Visa. Las tarjetas Mas se ven en menor medida al estar enfocadas hacia las compras en un solo negocio en particular. Las tarjetas Black y Platinum son pocas y se puede explicar por los requerimientos más altos de solidez económica necesarios para adquirirla, y el grueso de los clientes de la empresa se ubica en situaciones económicas más lejanas que las necesarias para estas tarjetas.

De la tabla además se puede concluir que existe similitud en este ámbito en los grupos de ambas iteraciones experimentales, de febrero y de marzo.

Con respecto a los cupos permitidos en las tarjetas de los clientes del experimento, estos varían y se puede apreciar su distribución en el Gráfico 3.

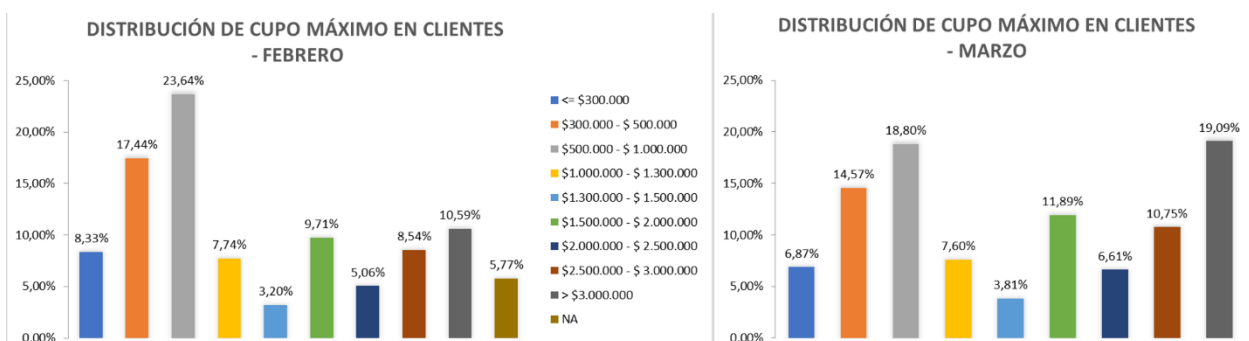


Gráfico 3: Distribución de los cupos máximos de las tarjetas de los clientes de los grupos experimentales de enero y febrero. El n total es de 40.000 en febrero y 45.000 en marzo.

Como se puede apreciar, hay una amplia variedad de intervalos de cupos de las tarjetas, y no hay amplias diferencias entre los grupos de ambos meses. El segmento de cupo más repetido es el cupo entre \$500.000 y \$1.000.000 en

ambos casos, seguido por la categoría inmediatamente menor. Esto nos indica que un poco menos de la mitad de los clientes se maneja con un cupo entre estos montos, por debajo del millón de pesos.

De este análisis se aprenden las tendencias más marcadas dentro de la muestra de individuos con la que se trabaja experimentalmente. También se concluyen las similitudes entre los grupos de febrero y de marzo que luego sustentan la decisión de juntar datos de ambos meses al momento de construir el modelo de uplift.

7. RESULTADOS EXPERIMENTALES

Al aplicarse los experimentos según las condiciones descritas en la sección 5 de este informe, se recopilan los resultados en el motor de base de datos de la empresa. Se recuerda que la variable de resultado del experimento es la contratación de un SAE (Súper Avance en Efectivo). En caso de que un cliente haya contratado el producto en el intervalo de tiempo entre el primer contacto publicitario del grupo expuesto (día 3 del mes correspondiente) y el final del mes, se considera una respuesta y se marca de manera correspondiente la columna VENTA_ATRIBUIDA con un 1 en la base de datos del proyecto.

Es importante explicitar que existen clientes seleccionados dentro del grupo expuesto a los que finalmente no se les llega a comunicar la campaña publicitaria. Las razones pueden ser varias. Una es que esos clientes pueden pertenecer a lo que se llama una purga en la empresa, que son clientes a los que se decide no comunicar por ver su situación de riesgo y/o crédito alterada. Otra razón para ser removido de la campaña es el hecho de haber contratado un SAE antes de haber recibido la publicidad. Otra más es que dicho cliente haya presentado un reclamo a Servicio al Cliente de la empresa o al Servicio Nacional del Consumidor después de ser escogido para el experimento.

Estas pérdidas ocurren en los experimentos de campo. El número real de clientes del grupo de tratamiento que efectivamente reciben las campañas publicitarias se denomina *n real*. El número propuesto al inicio en el diseño experimental se denomina *n propuesto*.

Dicho esto, los resultados de ambos experimentos se explicitan en la Tabla 4. Cabe recordar que el grupo de tratamiento o grupo expuesto se divide a su vez en grupo Expuesto Call-Center y grupo Expuesto Todo-canal, como se explicita en el diseño experimental de la sección 5 de este informe.

Grupo	FEBRERO				MARZO			
	n propuesto	n real	SAE	TR	n propuesto	n real	SAE	TR
Control	20.000	20.000	171	0,86%	15.000	15.000	142	0,95%
Expuesto	20.000	18.023	335	1,86%	30.000	27.558	499	1,81%
TOTAL	40.000	38.023	506	1,33%	45.000	42.558	641	1,51%
Expuesto Call-Center	10.000	9.060	152	1,68%	15.000	14.482	180	1,24%
Expuesto Todo-Canal	10.000	8.963	183	2,04%	15.000	13.076	319	2,44%
TOTAL	20.000	18.023	335	1,86%	30.000	27.558	499	1,81%

Tabla 4: Resultados experimentales de ambos experimentos. SAE = Número de productos SAE contratados del grupo correspondiente. TR = Tasa de Respuesta igual a (SAE/n real).

Según el cálculo del tamaño necesario explicitado en la sección 5.2, la diferencia en la Tasa de Respuesta (TR) utilizada para ese cálculo era de 0,4%. Según los resultados experimentales esa diferencia entre grupo de control y grupo de tratamiento es cercana a un 1% en ambos meses. Ya que el número utilizado en el cálculo marcaba una cota inferior para la validez del experimento, y en los resultados se superó ampliamente (por más del doble) se puede afirmar en concordancia con la teoría que ambos experimentos tienen la cantidad de individuos suficientes para ser estadísticamente significativos.

Al compararse con las campañas comunes de SAE de la empresa, la TR promedio ha sido ligeramente mayor, ya que ésta oscilaba entre el 0,9% y el 1,2%, mientras que el promedio total del experimento está cerca al 1,4%.

Otra observación es que en ambos meses se aprecia una mayor TR en el grupo expuesto a todos los canales que en el grupo expuesto a sólo llamadas, más acentuado en el mes de marzo.

Se procede a evaluar las diferencias entre los grupos, según los datos de la Tabla 4, en dos tests de student de una cola con una confianza del 95%. Tanto para el experimento de febrero como el de marzo, el p-valor del test del grupo expuesto contra el grupo de control es mucho menor a 0,001, por lo que, dada la confianza establecida del 95% para ambos tests, hay evidencia significativa para asegurar que existe una diferencia asociada al tratamiento entre los dos grupos.

Luego, evaluando las diferencias entre los grupos expuesto Call-center y expuesto Todo-canal, se puede observar que para los datos del mes de febrero el valor del p-valor es de 0,06 mientras que para marzo, el p-valor correspondiente es de un valor menor a 0,001. Nuevamente en ambos casos, (aunque en el de marzo la diferencia se hace mucho más evidente) el p-valor es menor al límite definido para rechazar la hipótesis nula de que no hay diferencia en las TR. Dado esto, la evidencia sugiere adoptar la hipótesis alternativa planteada en la sección 2.4, que las comunicaciones por todos los canales tienen mejores tasas de respuesta que las campañas difundidas solamente por call-center.

A continuación, se repiten los mismos grupos con tests de chi-cuadrado de confianza 95%, y se mantiene la misma conclusión.

Finalmente, hay un efecto que es posible observar y comparar que no está contemplado como un tratamiento definido en el diseño experimental. Al estar asignados los clientes a una campaña a través de call-center, puede pasar que se le llame pero no se le logre contactar, ya sea porque no conteste el teléfono, esté apagado, conteste otra persona, entre otros casos. Dado este hecho,

tanto para el grupo de sólo call-center como para el grupo multi-canal, hay clientes que no fueron publicitados a través de llamada.

Dado este hecho, se vuelve a comparar las tasas de respuesta de ambos grupos de tratamiento, pero esta vez considerando sólo a los clientes a los que efectivamente se había logrado contactar a través de la llamada, es decir, que contesten el teléfono y mantengan una conversación con el call-center.

Por otra parte, el hecho de que haya clientes del grupo Todo-canal, que no hayan sido contactados a través de llamada, significa que se forma a través de un experimento natural un grupo de clientes que fueron contactados a través de todos los canales menos el de call-center, al que se llama el grupo No-call.

Tanto los resultados de los grupos originales contando solamente las llamadas realizadas, como el grupo No-call, mantienen las conclusiones previas. Los resultados de comparar las tasas de respuesta de estos tres grupos son:

- Hay mejor tasa de respuesta en el grupo Todo-Canal que en el grupo sólo call-center.
- Hay mejor tasa de respuesta en el grupo Todo-Canal que en el grupo sólo No-Call.
- Hay mejor tasa de respuesta en el grupo Sólo-call que en el grupo sólo No-Call.

Estas 3 conclusiones fueron extraídas de tests de student y tests de chi-cuadrado con una confianza del 95%, y reafirman el supuesto de un principio de que el canal del call-center es más efectivo que el resto de canales, ya que la menor TR la tuvo el grupo que no fue contactado a través de call-center. Además, a pesar de haber controlado por llamadas efectivamente realizadas, se mantiene la mayor TR en el grupo Todo-canal que en el grupo Sólo-call, por lo que la evidencia sugiere adoptar la hipótesis alternativa planteada en la sección 2.4, de que las comunicaciones por todos los canales tienen mejores tasas de respuesta que las campañas difundidas solamente por call-center.

8. MODELO UPLIFT

Con los datos procesados como se ha descrito anteriormente, se procede a la fase de construcción del modelo de estimación de LIFT para los clientes.

La metodología seguida para el trabajo de datos y la creación del modelo es la de KDD, Knowledge Discovery in Databases, ampliamente utilizada en trabajos de estas características.

Para la construcción de los modelos de Uplift usando un algoritmo de Random Forest (como se describe en la sección 4.5), se utilizan funciones de la librería *Uplift* de R, creada por Guelman (2013).^[10]

8.1 Construcción

Para la construcción del modelo se trabajan en conjunto los datos de los experimentos de febrero y de marzo para generar un modelo combinado. Se definen:

- Variable dependiente y es VENTA_ATRIBUIDA (contratar SAE o no, binaria)
- Variable de tratamiento es GC (pertenecer al grupo de control o no, binaria)
- Covariables

Los pasos a seguir para construir un modelo son los siguientes:

1. Separar los datos de los clientes aleatoriamente en dos grupos, uno de entrenamiento y otro de testeo.
2. Entrenar el modelo usando la función *upliftRF*, que procede a estimar el LIFT para los clientes del grupo de entrenamiento.
3. Testear el modelo entrenado haciendo una predicción con la función *predict* y los datos del grupo de testeo.
4. Chequear el ajuste de la predicción usando matriz de confusión, cálculo de AUC y curva ROC.

Se iteró varias veces con distintas combinaciones de covariables buscando la significancia de las mismas y ajustar el modelo a la mayor y mejor información disponible. Las variables que se utilizaron en el modelo final se indican en la Tabla 5.

EMAIL	AVANCISTA	Dias_antiguedad	HABILITADO_COMPRA
SMS	MONTO_SIN_PREPAGO	TIENE_HIJO	HABILITADO_REPA
MKTD	CCSAE_COLOR	CASADO	DISP_AVANCE_TOTAL
CC	scoreCCSAE_MR	TIENE_BR	SEGMENTO_ACT_PPFF
FREQ	DIAS_MORA	REGION	SEGMENTO_ACT_RETAIL
L	DEUDA_TOTAL	TIENE_VEHICULO	SEGMENTO_ACT_CONVENIO
R	CUPO_TOTAL	RANGO_EDAD	SEGMENTO_ACT_SEGURO

Tabla 5: Lista de variables utilizadas para entrenamiento de modelo de Uplift Random Forest.

Las demás variables no se incluyeron debido a que no se tenía datos para todos los clientes, lo que no cumple con un requisito de la función *predict*. En siguientes iteraciones se agregarán las demás variables mencionadas en la Tabla 1.

Para definir el número de árboles a utilizar, se lleva a cabo una iteración entrenando modelos desde 10 árboles hasta un total de 1000, sumando de a 10 en 10 y evaluando el rendimiento a cada iteración. La conclusión es que el aumento del rendimiento empieza a ser despreciable alrededor de los 200 árboles, por lo que se fija el número a trabajar para el modelo en 200.

En cada nodo de un árbol se testean 5 variables, correspondientes al número recomendado por Guillén,^[11] que es la raíz cuadrada del número de columnas, redondeando hacia abajo. Al haber 28 variables en este modelo la raíz cuadrada sería 5,29, que al redondear entrega el número utilizado.

8.2 Evaluación de modelo

Un proceso usual de los procesos de entrenamiento de modelos de este estilo es la selección de variables en varias iteraciones. Una ventaja de usar un proceso basado en Random Forest es que el algoritmo mismo entrega la importancia de las variables utilizadas para entrenarlo, quitando esa parte del proceso.

Dado esto, se presenta la importancia de las variables de los modelos de febrero y marzo en la Tabla 6.

Ranking	Variable	Importancia relativa
1	REGION	15,87
2	scoreCCSAE_MR	13,56
3	DEUDA_TOTAL	11,87
4	RANGO_EDAD	10,95
5	Dias	9,7
6	DISP_AVANCE_TOTAL	9,13
7	MONTO_SIN_PREPAGO	8,51
8	CUPO_TOTAL	8,18
9	SEGMENTO_ACT_PPFF	2,93
10	SEGMENTO_ACT_CONVENIO	2,02
11	TIENE_BR	1,18
12	TIENE_VEHICULO	1,16
13	TIENE_HIJO	0,93
14	SEGMENTO_ACT_RETAIL	0,85
15	CASADO	0,8
16	L	0,59
17	SEGMENTO_ACT_SEGURO	0,53
18	CCSAE_COLOR	0,44
19	R	0,42
20	FREQ	0,25
21	AVANCISTA	0,09
22	DIAS_MORA	0,02
23	HABILITADO_REPA	0,00

Tabla 6: Importancia de las variables para el modelo entrenado.

Es interesante el hecho de que la variable que reporta una mayor importancia sea la región del cliente, ya que no se esperaría que fuera un influencia tan importante. En las siguientes variables se muestra un comportamiento más esperado, con variables que se considera muy relacionadas a un súper avance en efectivo, como la deuda, la antigüedad del cliente y el monto disponible. El caso de la variable scoreCCSAE_MR es especialmente interesante, ya que representa el SCORE que se ha mencionado bastante en este informe. Como se menciona en la sección 2.4, es común en empresas del rubro que se guíen con un modelo basado enteramente en este indicador para su priorización de intervenciones. En este modelo, tiene sentido que tenga una importancia alta, ya que significa en gran parte la probabilidad pura de que un cliente contrate un SAE, si bien en este modelo esta probabilidad se ve modificada y condicionada por la intervención en sí, que es el interés final de este proyecto.

Al hacer el análisis por separado entrenando un modelo para febrero y para marzo, se pudo observar que, si bien hay variaciones entre ambos meses, no hay cambios drásticos entre uno y otro. Una variable que es poco importante en febrero no pasa a ser muy importante en marzo, ni viceversa en ambos casos, hecho que impulsa la decisión de juntar ambos meses en un solo modelo. Se mantiene esa consistencia en los modelos de ambos meses.

Luego, se procede a evaluar el rendimiento de la predicción del modelo con los datos de testeo usando la función *performance* de la misma librería utilizada en esta sección. Esta función procede a hacer lo siguiente:

1. Calcula la diferencia entre las probabilidades estimadas por el modelo siguiendo la fórmula de LIFT.
2. Rankea las diferencias y las agrupa en 10 deciles ordenados de mayor a menor con igual número de individuos. Se asume que dentro de cada decil cada cliente tendrá la misma probabilidad de compra.
3. Calcula la diferencia real de las TR entre los grupos de tratamiento y control para cada decil, y lo guarda con el nombre de uplift.

Luego, se puede obtener el valor del uplift calculado para cada decil, con lo que se puede empezar el análisis que permitirá categorizar estos deciles en las 4 categorías descritas en la sección 4.1 e ilustradas en la . Para este fin se busca construir 3 herramientas, los uplift bins, la curva de uplift y la curva de Qini, como explica Karlsson (2019).^[12]

Primero, para tener una primera noción de los valores de LIFT, se busca construir los Uplift bins o baldes de Uplift. Para esto se ordena a los clientes en orden descendente según su LIFT y se les divide en deciles del mismo número de individuos. Luego se asumirá que todos los clientes del grupo tienen la misma probabilidad calculada por el LIFT, que corresponde al promedio del LIFT individual de cada uno. Finalmente se grafica como se observa en el Gráfico 4, para observar en una vista rápida la distribución del LIFT. También se presentan datos acerca de estos uplift bins en el Anexo 1.

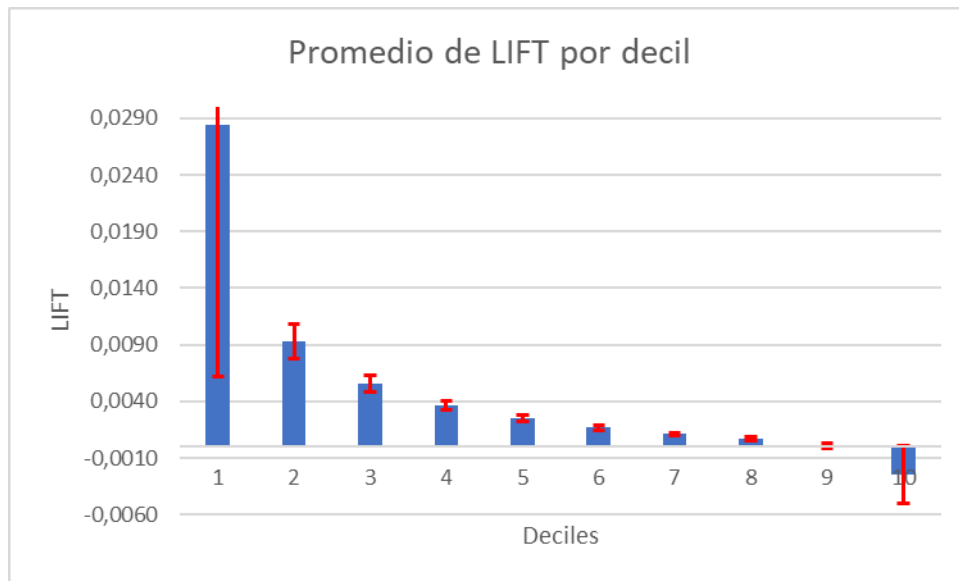


Gráfico 4: Baldes de uplift para LIFT calculado a partir del experimento. En rojo la suma y resta de la desviación estándar.

A partir de los baldes de uplift se observa a primera vista que los clientes con uplift positivo en promedio son más que los que tienen uplift negativo, es decir, hay más clientes influenciables que clientes perros dormidos que se molestan con la publicidad.

Luego, se procede a construir la curva de uplift, que grafica las ganancias acumuladas de contrataciones en razón de a qué porcentaje de la población se selecciona para la campaña de marketing, en el caso de seguir un sistema de priorización estricta según ranking de uplift.

Para esto se pueden utilizar los deciles calculados anteriormente, o construir un ranking a nivel de cliente individual, ordenándolos de forma decreciente según su valor de LIFT, e ir sumando las compras reales.

Con esta curva se pueden evaluar posibles decisiones de a cuánto porcentaje de la población seleccionar para una campaña (en el caso de priorizar a los individuos por uplift) y obtener resultados teóricos de las respuestas.

Otra manera de evaluar el rendimiento de un modelo de uplift es a través de la curva de Qini, similar a la curva de uplift pero con una diferencia fundamental. La curva de Qini grafica la diferencia en contrataciones incrementales entre el grupo de control y el grupo de tratamiento (a diferencia de la curva de uplift que considera ganancias acumuladas), calculadas por decil. Un ejemplo de cómo calcular los puntos de ambas curvas se puede observar en la Tabla 7, en donde las ganancias acumuladas (de la curva de uplift) se ven en la columna "Compras acumuladas", mientras que las

ganancias incrementales (de la curva Qini) se ven en la columna “Compras incrementales”.

Decil	Tratamiento		Control		Tratamiento - Control	
	Compras acumuladas	Clientes seleccionados	Compras acumuladas	Clientes seleccionados	Compras acumuladas	Compras incrementales
1 (10%)	11	1000	3	1000	8	8
2 (20%)	30	2000	10	2000	28	20
3 (30%)	44	3000	22	3000	50	22
4 (40%)	52	4000	33	4000	69	19
...

Tabla 7: Tabla ejemplo de cálculo de compras incrementales para curva de Qini. En este ejemplo la población total es de 10.000 para cada grupo, para simplificar el cálculo.

La curva de Qini, al igual que la de uplift, como se había descrito, puede ser usada para evaluar decisiones de a cuánto porcentaje de la población seleccionar para una campaña. Es común graficar ambas curvas junto a una curva que representa el resultado de una asignación completamente aleatoria a las campañas, que permite distinguir qué tanta diferencia hay entre hacer una priorización según uplift y una aleatoria.

Hacer esta evaluación usando la curva de Qini por sobre la curva de uplift tiene la ventaja de que, al tomar las ganancias incrementales en vez de las acumuladas, se captura el posible efecto negativo que puede tener el tratamiento en clientes poco receptivos (perros dormidos).

A continuación, en el Gráfico 5, se presenta la curva de Qini del modelo para los datos agregados de los experimentos de febrero y marzo.

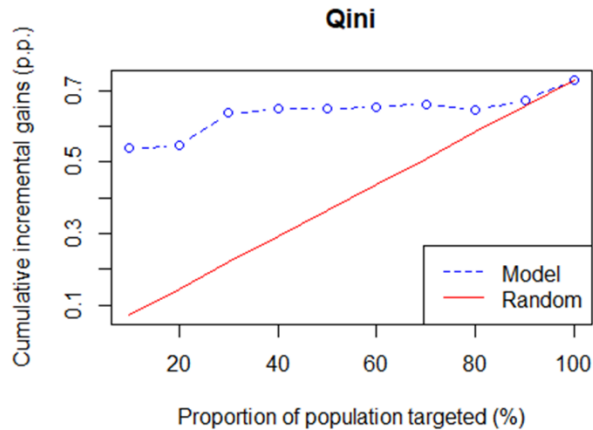


Gráfico 5: Curva de Qini del modelo entrenado. En el eje x se muestra el porcentaje de población seleccionado para una campaña hipotética, y en el eje y las ganancias incrementales en puntos porcentuales. La línea roja representa una selección aleatoria y la punteada azul una selección siguiendo el sistema priorizado por el modelo.

Se puede apreciar que gráficamente la curva de modelo se mantiene al avanzar el porcentaje de la población contactado, y que al final al hacerle publicidad al 100% de la muestra la ganancia incremental aumenta. Esto indica en una primera instancia que la ganancia de hacer llegar publicidad siempre va a ser mejor que no hacerlo, a pesar del efecto negativo que puede tener hacerle llegar publicidad a los clientes perros dormidos, ya que su número es reducido.

Sin embargo, al hacer un análisis más profundo, se observa que al apuntar las campañas a porcentajes pequeños de la población el incremental del sistema de LIFT para el grupo expuesto es mucho mayor que un sistema aleatorio de asignación. Mientras menos población se elija para hacerle llegar campañas, el sistema de LIFT mantiene un mejor rendimiento con respecto a un sistema de elección aleatorio.

Luego se puede observar también que en varios segmentos de la curva, el efecto incremental se mantiene cercano a ser constante, por ejemplo, entre el 40% y el 70% de la población el incremental es muy pequeño, lo que indica el efecto de los clientes causa perdida y los clientes asegurados, además del efecto negativo de los clientes perros dormidos que contrarrestan el incremental de los clientes influenciables de estos deciles. Dado esto, en teoría, se obtendría el mismo resultado en la campaña publicitando al 40% de la población que al 70% de la misma, lo que supone un ahorro de costos de un 43% si se asume un costo no-variables para todos los clientes.

La conclusión que se toma en este caso es que dada la composición de los valores de LIFT de los clientes, siempre se tendrá un resultado en tasa de respuesta mayor o igual al aumentar el número de clientes contactados, pero hay rangos en donde se obtiene el mismo resultado al contactar a un número mayor de clientes hasta cierto punto, por lo que lo óptimo tomando en cuenta los costos no siempre es publicitar a más clientes, sobre todo si no se puede publicitar a todos.

8.3 Construcción de modelo específico para canales

Luego, se replica el procedimiento pero entrenando un segundo modelo exclusivamente para los distintos tratamientos aplicados. Se separa para este fin solamente a los clientes que estuvieron expuestos a alguno de los dos tratamientos, excluyendo al grupo de control, y luego se toma:

- Variable dependiente y es VENTA_ATRIBUIDA (contratar SAE o no, binaria)
- Variable de tratamiento es EXPUESTO_CALL (pertenecer al grupo de sólo call-center o al grupo de Todo-canal, binaria)
- Covariables

Este modelo estaría analizando entonces el efecto de hacerles llegar campañas a través de otros canales (aparte de llamada) a los clientes versus el no hacerlo, ya que, de esta base, todos los clientes están clasificados para ser contactados por call-center.

Se procede a construir los baldes de canales para este modelo como se puede ver en el Gráfico 6.

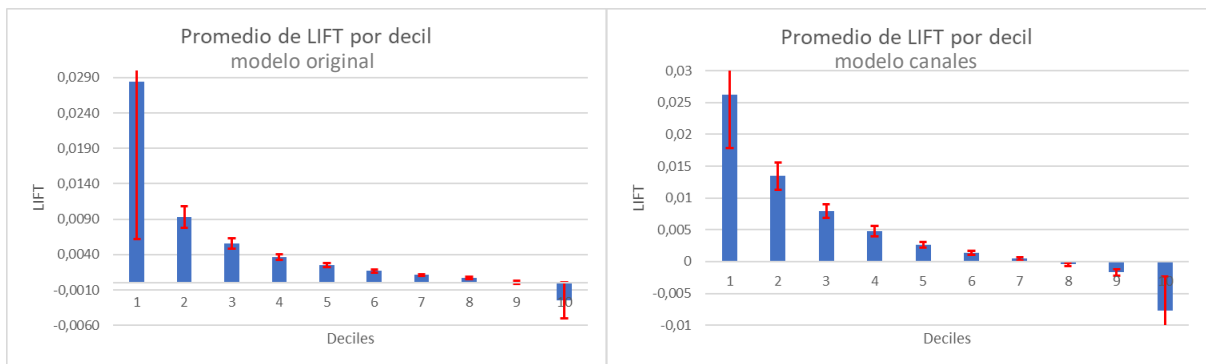


Gráfico 6: Baldes de uplift para LIFT calculado con ambos modelos. A la izquierda el Gráfico 4 y a la derecha los baldes para el modelo de canales recién descrito. En rojo la suma y resta de la desviación estándar.

A primera vista, al comparar los uplift bins con el modelo anterior, se puede ver que hay más clientes en el lado negativo del LIFT, en los bins 8, 9 y 10, cuando en los bins anteriores sólo había LIFT negativo en el último decil. Esta noción se ve reforzada al analizar las estadísticas en la Tabla 8, donde se puede ver también que el máximo valor de LIFT es menor al anterior y el mínimo es más negativo que el anterior. En la tabla y en el gráfico se puede notar que los primeros 4 deciles de este modelo alcanzan más alto que los del anterior, pero que después de esto descienden los valores de LIFT más que en el anterior.

Decil	Promedio	Máximo	Mínimo	Mediana	Desviación estándar
1	0,0263	0,0660	0,0176	0,0236	0,0084
2	0,0135	0,0176	0,0101	0,0133	0,0022
3	0,0080	0,0101	0,0062	0,0078	0,0011
4	0,0048	0,0062	0,0035	0,0048	0,0008
5	0,0027	0,0035	0,0019	0,0026	0,0005
6	0,0014	0,0019	0,0009	0,0014	0,0003
7	0,0005	0,0009	0,0000	0,0005	0,0002
8	-0,0004	0,0000	-0,0009	-0,0004	0,0003
9	-0,0017	-0,0009	-0,0027	-0,0016	0,0005
10	-0,0077	-0,0027	-0,0369	-0,0055	0,0054

Tabla 8: Estadísticas de LIFT por decil de uplift bins para modelo de canales.

Esto indica que los clientes son más sensibles a la publicidad cuando se refiere a la decisión de recibir campañas a través de múltiples canales; los más influenciados en general reaccionan mejor, pero hay más clientes que se sienten molestados y se desinteresan en adquirir un producto. De esto se desprende que existe un riesgo mayor de molestar a un cliente y generar un efecto negativo al agregar publicidad multi-canal a la comunicación por llamada, a pesar de sumar mayores conversiones. Dado esto, aumenta también la necesidad de conocer el tipo de cliente y su probable reacción a las intervenciones.

Esta conclusión queda más clara al mirar la curva de Qini. En ella se observa, en el Gráfico 7, que hay una diferencia en la curva del modelo, que no es ascendente, sino que tiene un peak de ganancia incremental, y luego desciende.

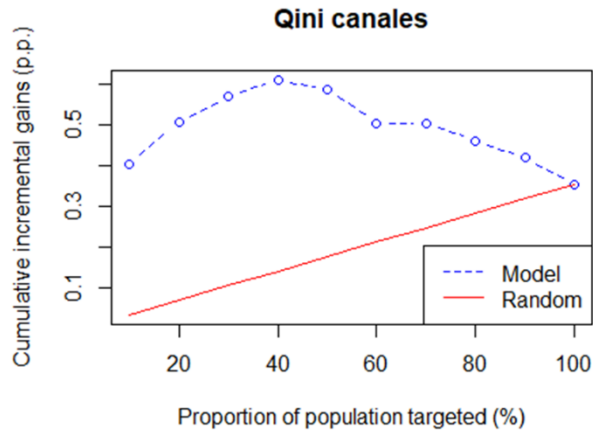


Gráfico 7: Curva de Qini para el modelo de comparación de tratamientos Todo-canal v/s Sólo-call. En el eje x se muestra el porcentaje de población seleccionado para una campaña hipotética, y en el eje y las ganancias incrementales en puntos porcentuales. La línea roja representa una selección aleatoria y la punteada azul una selección siguiendo el sistema priorizado por el modelo.

Esto se debe a que en esta comparación el número de clientes con un valor de LIFT negativo es mayor, hay más clientes *sleeping dogs*. El resultado es que, a diferencia del caso anterior, no siempre el resultado será igual o mejor al aumentar la cantidad de clientes contactados, sino que la mejor tasa de respuesta (0,6 puntos porcentuales por sobre el grupo de control) se obtendrá al hacerle llegar campañas a un 40% de los clientes en el sistema ordenado por LIFT. Al contactar a más de eso el incremental empieza a bajar, es decir que empieza a haber menos porque se activa el efecto de los clientes que son susceptibles de forma negativa a la publicidad.

La conclusión que se toma es que no siempre es mejor contactar a través de todos los canales a los clientes una vez que se decide contactarles a través de call-center, ya que el riesgo de desincentivar la conversión es mayor que el posible beneficio esperado. A pesar de que los promedios puedan indicar lo contrario, hay algunos comportamientos indeseables (contactar y molestar a clientes *sleeping dogs*) que se ocultan tras esos promedios que sustentan la conclusión.

Luego, tomando al grupo No-call, que corresponde al grupo de clientes que se habían designado en el grupo Todo-canal, pero que finalmente no fueron contactados por llamada (como se describe en la sección **Error! Reference source not found.**), se puede comparar el efecto de agregar la llamada a un grupo de clientes que se decide contactar por los otros canales, e-mail, SMS y carta, de manera análoga a la recién descrita.

Esta vez, los resultados se asemejan más al modelo original. Si se analiza la curva de Qini de este modelo en el Gráfico 8, se observa que en la mayoría de su longitud la curva tiene variaciones pequeñas de incremental. La diferencia principal con el modelo original es que contactar al 100% de la población entrega un incremental esperado peor que si se contactara a menos clientes, un 90%, 80% u otros cortes.

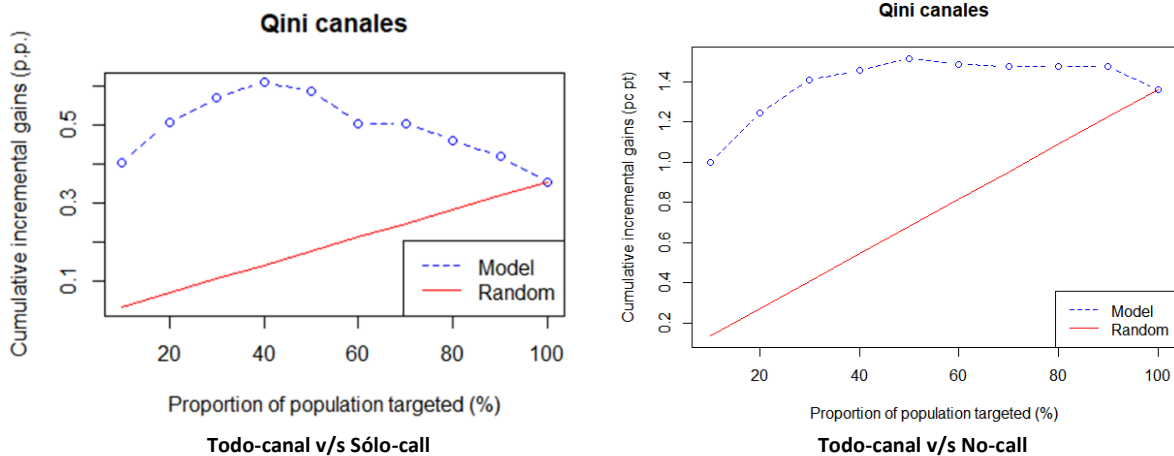


Gráfico 8: Curvas de Qini para los modelos de comparación de tratamientos Todo-canal v/s Sólo-call (Gráfico 7) y Todo-canal v/s No-call. En el eje x se muestra el porcentaje de población seleccionado para una campaña hipotética, y en el eje y las ganancias incrementales en puntos porcentuales. La línea roja representa una selección aleatoria y la punteada azul una selección siguiendo el sistema priorizado por el modelo.

Esto indica que, al igual que en el análisis anterior del modelo de Todo-canal v/s Sólo-call, hay un número de clientes *perros dormidos* que influyen en el resultado de la campaña, si bien es una influencia menor que en el caso anterior.

En este caso entonces, al tener un número de clientes que se van a contactar a través de e-mail, SMS y/o carta física, no siempre entrega un mejor resultado incremental llamarles a través de call-center además, sino que el mejor resultado esperado se encontraría al contactar a un 50% de ellos, aunque el incremental que se pierde al contactar más que eso no es tanto como en el caso anterior, siempre que se contacte a menos del 100% de los clientes.

El resultado más positivo de la curva de Qini en este caso se debe al hecho de que el canal más efectivo para las campañas está demostrado que es la llamada por call-center, y se está evaluando justamente el efecto de agregar una llamada por call-center a una publicidad de otros canales. Esto explica también los valores más elevados de respuesta incremental, que llegan a estar

sobre 1,4 puntos porcentuales, a diferencia del otro modelo en donde se alcanzaban cifras menores.

Habiendo analizado ambos casos, se concluye que la estrategia de contactar a los clientes a través de call-center y además a través de los demás canales es efectiva para los clientes más influenciables, con más valor de LIFT, pero no es la mejor estrategia para todos los clientes, sobre todo los *sleeping dogs*. Es decir, no siempre es mejor contactar a todos los clientes a través de todos los medios disponibles, sino que depende de cada cliente, lo que refuerza la segmentación propuesta en la sección 4.1 del informe.

Por último, si bien el valor de LIFT varía para los clientes según ambos modelos de canales, en la mayoría de los casos los clientes se mantienen en su decil o a un decil de distancia del decil en el que habían sido clasificados en el modelo original, por lo que se puede usar como referencia el decil de clasificación del modelo original, con el que se seguirá evaluando, como un indicador de receptividad a campañas a través de múltiples canales a través de call-center.

8.4 Aplicación de modelo a campaña real

Se desea evaluar si el modelo es generalizable y observar si efectivamente tiene capacidad de pronóstico fuera de la muestra experimental. Para esto, se propone la siguiente aproximación:

- 1) Tomar los datos necesarios para estimar LIFT de clientes que hayan participado en una campaña en un mes distinto.
- 2) Estimar su LIFT con el modelo.
- 3) Construir curva de Qini.
- 4) Comparar los resultados estimados de la curva de Qini con los resultados reales y evaluar el rendimiento del modelo con respecto al sistema actual de priorización.

Sin embargo, para poder cumplir con el cuarto paso y comparar los resultados teóricos del modelo con los reales se necesita un grupo de control para comparar los puntos incrementales. Dado esto, se toman datos de campañas del mes de enero, ya que en ese mes la empresa apartó un grupo de control de clientes seleccionados aleatoriamente ese mes para que no fueran contactados por las campañas de SAE, justamente para que funcionaran como grupo de control de las campañas.

Entonces, se procede a tomar los datos de los clientes que participaron de las campañas de SAE en enero y los de este grupo de control, y se estima su LIFT utilizando el modelo según sus covariables.

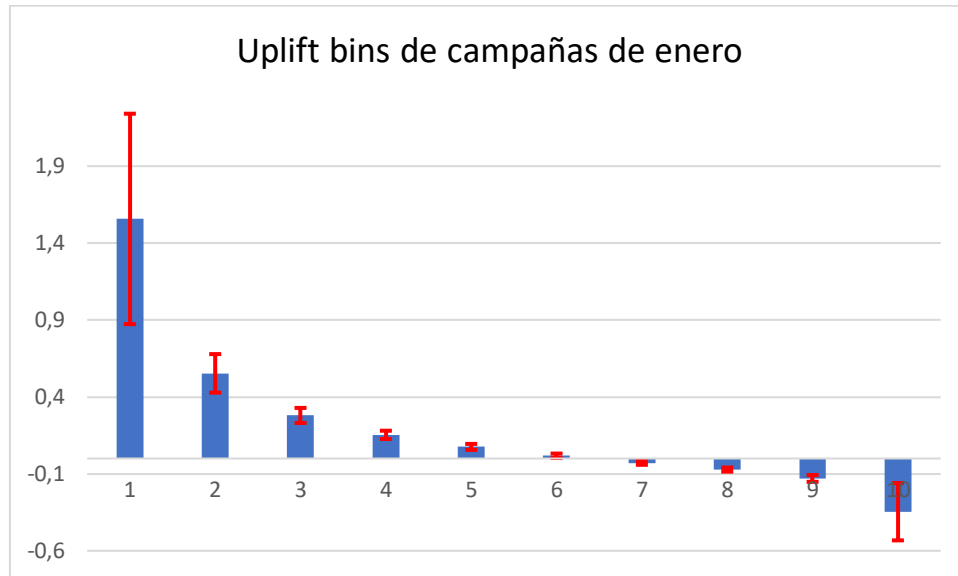


Gráfico 9: Baldes de uplift para LIFT calculado en clientes de campañas de enero. En rojo la suma y resta de la desviación estándar.

La curva de Qini del resultado se puede observar en el Gráfico 10.

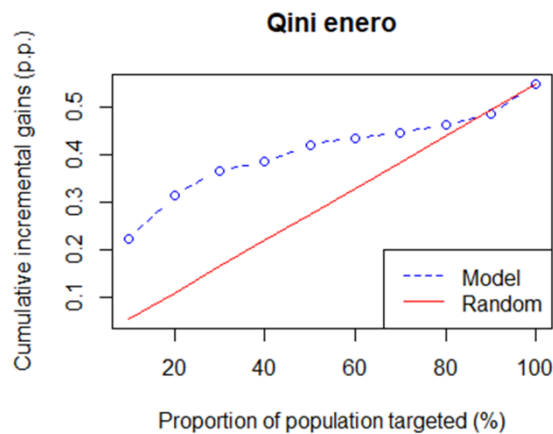


Gráfico 10: Curva de Qini de clientes de campañas reales de enero. En el eje x se muestra el porcentaje de población seleccionado para una campaña hipotética, y en el eje y las ganancias incrementales en puntos porcentuales. La línea roja representa una selección aleatoria y la punteada azul una selección siguiendo el sistema priorizado por el modelo.

La curva de Qini se asemeja a la curva de los experimentos, en el Gráfico 5, por lo que se comprueba el ajuste del modelo. El siguiente paso es evaluar su rendimiento, por lo que se ordenan los clientes según LIFT y se evalúan los resultados de hacerle llegar campañas de publicidad a distintos porcentajes de la población. Luego, se realiza el mismo ejercicio pero ordenando los clientes de manera aleatoria. Finalmente, con el fin de comparar el sistema de priorización por LIFT con el sistema actual de la empresa, se ordenan los clientes según su scoring de propensión, y se asume la misma estrategia de publicitar a distintos deciles de la población. Al ser 773.306 los clientes totales de la muestra, publicitar al primer decil correspondería a contactar a 77.331, contactar a los dos primeros deciles, equivalente al 20% de la población, sería contactar a 154.661, y así. Los resultados se pueden observar en la Tabla 9.

Porcentaje	TR incremental		
	Uplift	Aleatorio	Scoring
10%	0,09%	-0,74%	-0,08%
20%	0,41%	-0,59%	0,37%
30%	0,59%	-0,47%	0,57%
40%	0,70%	-0,33%	0,66%
50%	0,73%	-0,17%	0,68%
60%	0,79%	-0,02%	0,69%
70%	0,80%	0,13%	0,61%
80%	0,72%	0,24%	0,43%
90%	0,55%	0,18%	0,04%
100%	-	-	-

Tabla 9: Comparación de tasas de respuesta tomando distintos porcentajes de población ordenada según los tres sistemas distintos. Todas las TR se calculan sobre el n total de clientes (773.303).

Como se puede observar, la mayor tasa de respuesta incremental se encuentra en el sistema de ordenamiento según Uplift, lo que apoya la idea de que usar este sistema podría obtener mejores resultados que un sistema aleatorio y que el sistema actual de la empresa.

Por otra parte, se puede considerar el efecto de los clientes *sleeping dogs* que se caracterizan por ser clientes con LIFT negativo, que tienen más probabilidades de comprar el producto si no son contactados que si lo son. En los datos reales hay un gran número de clientes a los que se les estima como *sleeping dogs*, con uplift negativo, pero a los que efectivamente sí se les hizo llegar campañas, por lo que no se estaría contando en los datos reales el efecto positivo de esas ventas que se ganan al no hacerles llegar publicidad a los *sleeping dogs* que estén en ese 60% de población no contactada.

Para estimar ese efecto se estimó una *TRSD*, en relación a *Tasa de Respuesta Sleeping Dogs*, que se estimó tomando la tasa de respuesta de los clientes del grupo de control que tienen LIFT negativo, como se describe en la Ecuación 5. Luego, esa *TRSD* funciona como ponderador para el número de sleeping dogs no contactados en cada estrategia.

$$TRSD = \frac{N^{\circ} \text{ de clientes del grupo de control con LIFT} < 0 \text{ que compraron SAE}}{N^{\circ} \text{ de clientes del grupo de control con LIFT} < 0} = 0,31\%$$

Ecuación 5: Cálculo de Tasa de Respuesta Sleeping Dogs

Luego, si por ejemplo en el sistema de uplift se hubiera dejado tranquilos sin publicidad a 100 sleeping dogs, el efecto en la tasa de respuesta sería como se observa en la Ecuación 6.

$$\text{Efecto de Sleeping Dogs en la estrategia} = \frac{0,31\% * 100}{n \text{ total de clientes}} = \frac{31}{773.303}$$

Ecuación 6: Estimación del efecto sleeping dogs en la tasa de respuesta.

Luego, ese efecto Sleeping Dogs se puede agregar a la Tasa de Respuesta incremental calculada anteriormente para obtener un estimado de los resultados con este efecto. Al haber 320.738 clientes en la muestra con un valor de LIFT negativo en los datos, los resultados se observan en la Tabla 10.

Porcentaje	TR incremental			TR incremental + efecto SD		
	Uplift	Aleatorio	Scoring	Uplift	Aleatorio	Scoring
10%	0,09%	-0,74%	-0,08%	0,21%	-0,62%	0,04%
20%	0,41%	-0,59%	0,37%	0,54%	-0,49%	0,48%
30%	0,59%	-0,47%	0,57%	0,71%	-0,38%	0,67%
40%	0,70%	-0,33%	0,66%	0,82%	-0,25%	0,76%
50%	0,73%	-0,17%	0,68%	0,86%	-0,1%	0,77%
60%	0,79%	-0,02%	0,69%	0,92%	0,03%	0,77%
70%	0,80%	0,13%	0,61%	0,89%	0,16%	0,67%
80%	0,72%	0,24%	0,43%	0,78%	0,27%	0,47%
90%	0,55%	0,18%	0,04%	0,58%	0,20%	0,06%
100%	-	-	-	-	-	-

Tabla 10: Comparación de resultados de cada sistema de priorización de campañas según porcentaje de la población contactado. Todas las TR se calculan sobre el n total de clientes (773.303).

Se puede observar que la mayor TR incremental total sigue estando en el sistema de Uplift, que además es el sistema que más ganancia obtiene del hecho de no contactar a los clientes *sleeping dogs*, ya que no contacta a ninguno, lógicamente, ya que el sistema se basa en ese ranking. Se pueden observar estos resultados en una curva en el Gráfico 11.

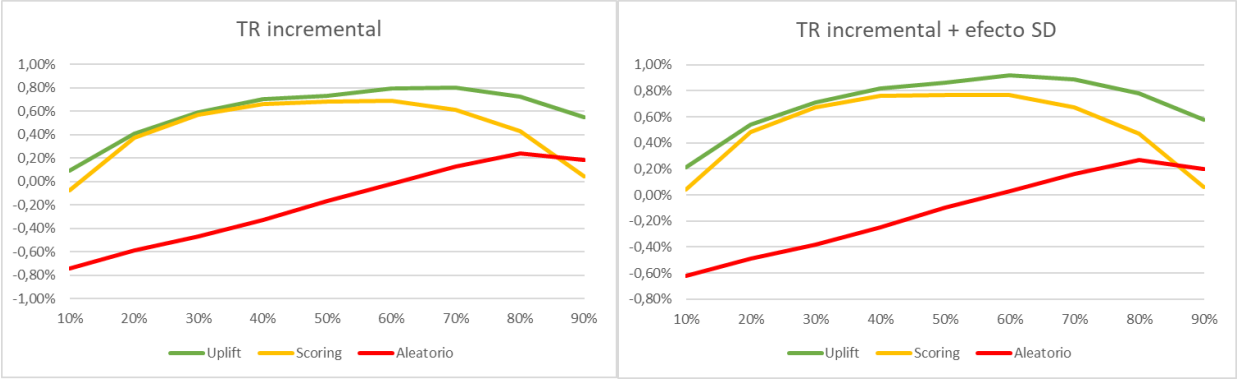


Gráfico 11: Gráfico de resultados de sistemas de priorización según porcentaje de población contactado.

Los resultados muestran que la conclusión se cumple sin importar el porcentaje de la población que se tome, el sistema de priorización por Uplift siempre tiene mejores resultados incrementales.

Para comprobar que el modelo y las conclusiones fueran aplicables a cualquier período de tiempo, se repitió el procedimiento entero para las campañas reales del mes de diciembre, que también contó con grupo de control, y los resultados se mantuvieron dentro de los mismos rangos, siempre estando el sistema de Uplift por sobre los otros dos en los resultados incrementales. También se mantuvieron en el mismo rango, con diferencias despreciables las estimaciones del valor de LIFT, por lo que el modelo es robusto en comparación a campañas de distintos periodos.

9. ANÁLISIS DE RESULTADOS

Ya comprobado que el modelo puede ser aplicado y obtener mejores resultados esperados en Tasa de Respuesta que un sistema aleatorio y que el sistema actual de la empresa, se analizan las relaciones que pueda tener el LIFT con otros aspectos importantes al momento de escoger qué clientes reciben campañas de marketing.

9.1 Comportamiento de pago de clientes

Ya que el proyecto se aplica a productos financieros, más específicamente a un avance de dinero, los clientes que toman el producto quedan en deuda con la empresa, y deben pagar su deuda mensualmente en los meses posteriores. Los clientes que pagan regularmente son los mejores desde el punto de vista de la empresa, mientras que los clientes morosos son menos deseables, *ceteris paribus*.

Se busca estudiar si hay una relación directa entre el valor de LIFT y el comportamiento de pago de los clientes, por lo que se tomó la información de pago de deuda de éstos en los 4 meses posteriores a los experimentos (abril, mayo, junio y julio) para llevar a cabo este análisis. Para fines del estudio se categorizó a los clientes que:

1. Tomaron un SAE en alguno de los experimentos de febrero o marzo.
2. Tienen deuda que pagar en los 4 meses posteriores.
3. Registran un pago de deuda en esos 4 meses.

De cumplir con las 3 condiciones se clasifica al cliente como que pagó deuda, mientras que si cumple con las primeras 2 condiciones, pero no paga deuda en alguno de los 4 meses del estudio, se considera como no-pago. Cualquier cliente que no cumpla con las primeras dos condiciones se excluye de este estudio. Luego se puede calcular una tasa de pago de deudas considerando el número de clientes que pagan dividido por el número de clientes totales, que alcanza un número de 91%.

Para relacionarlo con el LIFT, se toman los deciles calculados en los experimentos de febrero y marzo y se analiza la distribución de estos clientes en los deciles, como se puede observar en la Tabla 11.

Decil de LIFT	Cientes totales con deuda	Cientes que pagan deuda	Cientes que no pagan deuda	Porcentaje de pago
1	101	98	3	97%
2	100	96	4	96%
3	100	83	17	83%
4	101	78	23	77%
5	100	99	1	99%
6	100	96	4	96%
7	101	94	7	93%
8	100	91	9	91%
9	100	86	14	86%
10	100	91	9	91%

Tabla 11: Comportamiento de pago de deuda de clientes según decil de LIFT.

De estos resultados no se puede afirmar que haya una relación directa a simple vista entre el porcentaje de pago de deuda con el LIFT, ya que la distribución se desordena en los deciles, no sigue un orden establecido. Se esperaría que si hubiera una correlación positiva todos los porcentajes más altos estuvieran en los primeros deciles y los peores en los últimos, y viceversa para una correlación negativa, pero no es el caso en ninguna de las dos opciones. Por otra parte, el número reducido de clientes en cada decil no es suficiente como para poder afirmar una relación con el poder suficiente, por lo que se concluye que de esta manera no se puede afirmar que haya una relación directa entre el LIFT de los clientes y su comportamiento de pago. Dado esto, es necesario obtener otro constructo para medir el valor del cliente, que es el parámetro que es interesante obtener para poder realizar el análisis entre LIFT y valor de cliente para poder diseñar sistemas de priorización de clientes.

Teniendo en cuenta esto, se propone estimar la probabilidad de un cliente de pagar su deuda usando un modelo de regresión logística, que considere solamente variables relacionadas directamente con la tarjeta del cliente. Como variable a predecir se utiliza la misma clasificación de pago o no-pago anterior.

Las variables utilizadas para predecir el comportamiento de pago se muestran en la Tabla 12. Para el entrenamiento del modelo se utiliza un 80% de la población de los experimentos, y para su testeo se utiliza el 20% restante.

AVANCISTA	scoreCCSAE_MR	HABILITADO_COMPRA
MONTO_SIN_PREPAGO	CUPO_TOTAL	HABILITADO_REPA
CCSAE_COLOR_NUEVO	DIAS_MORA	TCA

Tabla 12: Variables de entrenamiento de modelo de predicción de probabilidad de pago.

Los coeficientes de cada variable se adjuntan en el Anexo 1, en el que se puede ver que todas las variables menos una tienen una significancia mayor al 99%. También se adjuntan en los Anexos 2 y 3 las matrices de confusión de los datos de entrenamiento y testeo. La predicción al evaluarla con los datos reales entrega un accuracy de un 75% tanto para los datos de entrenamiento como para los datos de testeo, y el área bajo la curva de ROC es de 0,8 para ambos casos como se puede apreciar en el Gráfico 12, por lo que es seguro concluir que tiene suficiente poder explicativo, se considera como un buen predictor, y que no se está sobreajustando.

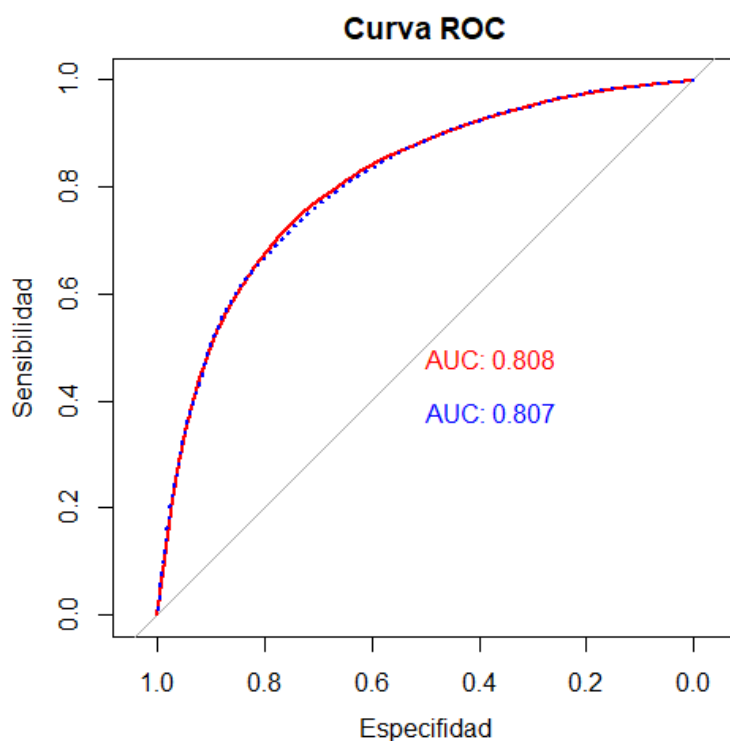


Gráfico 12: Curva ROC de modelo de predicción de pago de deuda. La línea roja continua representa el ROC de los datos de entrenamiento y la línea azul punteada la de los datos de testeo.

Como resultado del modelo, se tiene ahora una probabilidad de pago de deuda para cada cliente, por lo que se procede a estudiar la relación entre esta probabilidad y el valor de LIFT de cada cliente, de manera similar a la anterior, agrupando por deciles de LIFT y analizando el promedio de la probabilidad de pago. Los datos se pueden ver en la Tabla 13, en donde se muestran por decil las probabilidades de pago promedio estimadas por decil, y el número estimado de clientes que pagan la deuda por decil, considerando en estos

últimos como clientes que pagan a todos los que tengan probabilidad de pago mayor a un 50%.

Decil de LIFT	Clientes en decil	Probabilidad promedio de pago de deuda	Nº estimado de clientes que pagan	Tasa de pago
1	8267	83%	8.112	98%
2	8267	76%	7.701	93%
3	8267	69%	6.795	82%
4	8267	62%	5.930	72%
5	8267	59%	5.507	67%
6	8268	55%	4.784	58%
7	8267	47%	3.556	43%
8	8267	45%	3.289	40%
9	8267	48%	3.743	45%
10	8267	65%	5.997	73%

Tabla 13: Resultados de modelo de predicción de pago de deuda por decil de LIFT. Tasa de pago se calcula como número de clientes que pagan la deuda sobre el número de clientes del decil.

En esta tabla sí se observa una tendencia más marcada en relación a los deciles de LIFT, teniendo los deciles más altos la probabilidad promedio de pago y la tasa de pago más altas, junto con el último decil. Esto estaría indicando que los clientes que más propensos a pagar sus deudas, y por ende los que más atractivos le parecen a la empresa, serían los clientes clasificados como los más influenciables y los clientes *sleeping dogs* más extremos. Esto refuerza la idea de priorizar las campañas de productos financieros para asegurarse de que les lleguen a los clientes influenciables, que son los que mejor probabilidad tienen de contratar un producto si son contactados, y además son los que más probabilidades en promedio tienen de pagar su deuda con la empresa. Por otro lado, también se acentúa la idea de asegurarse de que no les lleguen campañas a los clientes más sensibles a molestarse por la publicidad, ya que de esta manera se maximiza la probabilidad de que tomen un producto financiero, y tienen una probabilidad promedio aceptable (la cuarta mejor de los deciles) de efectivamente pagar sus deudas. Entonces, la recomendación a partir de este análisis sería guiarse por el LIFT al priorizar clientes para campañas y la conclusión sería que no hay una diferencia entre un sistema de priorización de LIFT puro y uno según un *Value-Uplift* calculado de esta manera.

9.2 Rentabilidad de clientes

Otra variable de interés con la que se puede analizar la relación con el LIFT es la rentabilidad de cada cliente. Si se deseara estimar el valor de un cliente en base a su rentabilidad, es interesante estudiar si al hacer un sistema de priorización considerando LIFT, se debe modificar la selección en base a qué tan rentable es un cliente.

La empresa cuenta con los datos de rentabilidad (entendida como margen, ingresos – costos) de cada cliente en cada mes, así que se procede a analizarlos, tomando en cuenta como rentabilidad del cliente la rentabilidad del mes anterior, es decir, la rentabilidad de diciembre para cada cliente analizado en el mes de enero. Para analizar la relación de la rentabilidad de cada cliente con el LIFT se sigue la misma estrategia que con el comportamiento de pago, con la salvedad de que se toma el dato de rentabilidad de todos los clientes, no sólo los que hayan tomado un SAE, ya que se tiene la estimación de LIFT de todos los clientes. Los resultados se ven en la Tabla 14.

Decil	Nº clientes	Rentabilidad de clientes		
		Promedio	Máxima	Mediana
1	8.500	\$8.785	\$857.726	\$9.805
2	8.500	\$12.187	\$1.128.295	\$6.196
3	8.500	\$8.106	\$1.163.822	\$3.386
4	8.500	\$5.143	\$1.075.351	\$1.851
5	8.500	\$4.060	\$673.221	\$1.200
6	8.500	\$3.351	\$989.387	\$329
7	8.500	\$2.154	\$532.344	-\$120
8	8.500	\$2.025	\$663.047	-\$17
9	8.500	\$3.189	\$455.864	\$380
10	8.500	\$10.113	\$834.996	\$2.471

Tabla 14: Análisis de rentabilidad de clientes por deciles de LIFT.

En esta tabla sí se puede apreciar una relación con los deciles de LIFT. Se observa que los mayores valores de rentabilidad están agrupados en los primeros 5 deciles y en el último. Esto quiere decir que en esos deciles se encuentran los clientes más rentables para la empresa. Entre el decil 6 y el 9 están los clientes que menos rentabilidad entregan. Si se hace la relación de estos deciles con las categorías presentadas en la sección 4.1 del documento, los clientes con más rentabilidad serían los más influenciados y los *sleeping dogs* más extremos, mientras que los que menos rentabilidad tienen serían los

clientes asegurados o causas perdidas, más los *sleeping dogs* menos marcados.

De esto se concluye que una estrategia en donde se pone énfasis en publicitarle a los clientes influenciables y en no publicitarle a los clientes que se puedan molestar por ello sería positivo, ya que se maximiza la probabilidad de que estos clientes contraten los productos y además, estos clientes son los que en promedio más rentabilidad le entregan a la empresa a través de esas compras.

Dada la conclusión anterior, del comportamiento de pago, tiene sentido esta conclusión de rentabilidad de clientes, ya que para la empresa es más rentable un cliente que paga sus deudas mes a mes que uno que entra en mora o es irregular con sus pagos.

9.3 Montos y costos

El producto del Súper Avance en Efectivo utilizado en este proyecto se caracteriza por tener diferencias entre un cliente y otro, siendo la más notable de ellas el monto solicitado por el cliente como préstamo a la empresa. Para la empresa, los préstamos de monto más alto son más atractivos que los montos pequeños, por lo que se desea analizar los montos de los SAE que contratan los clientes, y si estos montos podrían tener una influencia en la decisión de elegir el porcentaje de población al que dirigir una campaña.

Dado esto, se toman los datos de la campaña real de enero y se crea la Tabla 15 en la que se analizan los montos de los SAE tomados por los clientes de según a qué porcentaje de la población se le publicita, usando un orden de publicidad ordenado por LIFT, de manera acumulada. Para este análisis se toman en cuenta sólo los SAE contratados en la campaña real, sin efecto estimado de *sleeping dogs* o tasa incremental con respecto a un grupo de control. Luego, se estudian las posibles diferencias entre porcentaje de población a contactar con el fin de luego poder escoger límites recomendados de población a contactar.

Porcentaje de la población contactada	Monto acumulado SAE	Diferencia de monto con decil anterior	Diferencia de monto con decil anterior en porcentaje
10%	\$10.515	-	-
20%	\$14.341	\$3.826	27,37%
30%	\$16.603	\$2.262	16,18%
40%	\$17.994	\$1.392	9,96%
50%	\$19.043	\$1.049	7,51%
60%	\$19.982	\$938	6,71%
70%	\$20.611	\$630	4,51%
80%	\$21.189	\$578	4,14%
90%	\$22.052	\$863	6,17%
100%	\$24.492	\$2.440	17,46%

Tabla 15: Monto acumulado de ventas de SAE de campañas de enero según porcentaje de la población contactada, en millones de pesos.

Luego, se busca analizar la relación con los costos de hacerles llegar campañas a los distintos porcentajes de la población, por lo que se repite el ejercicio pero esta vez analizando los costos de publicidad reales de la campaña de enero. Para este fin es importante notar que se tomaron los costos reales de las campañas de la empresa, que corresponden a \$0,2 por email enviado, \$20 por SMS enviado, \$90 por carta física enviada y un 5% del monto del SAE en costos de call-center solamente en el caso de que la venta se haya concretado a través de ese canal. Recordar aquí que el costo diferenciado de las llamadas por call-center se debe a que los llamantes son externos a la empresa y el contrato así lo explicita. Los resultados se pueden ver en la Tabla 16.

Porcentaje de la población contactada	Costos de publicidad acumulados	Diferencia de costos con decil anterior	Diferencia de costos con decil anterior en porcentaje
10%	\$261	-	-
20%	\$356	\$95	25,98%
30%	\$414	\$58	15,93%
40%	\$451	\$36	9,91%
50%	\$480	\$30	8,14%
60%	\$504	\$23	6,35%
70%	\$523	\$20	5,36%
80%	\$538	\$15	4,15%
90%	\$562	\$24	6,58%
100%	\$627	\$64	17,61%

Tabla 16: Costos de campañas acumulados para campañas de SAE en enero según porcentaje de la población contactada, en millones de pesos.

Los costos directos de campaña son despreciables en relación a los montos de SAE, pero si analizamos el porcentaje de crecimiento al pasar de un decil a otro de población contactada, se puede ver que hasta el 40% de población contactada el monto crece más que los costos al pasar al siguiente decil, pero en el 50% se revierte esa relación. Luego en el 60% vuelven a crecer los montos más que los costos y de ahí para abajo siempre que se suma un 10% más de población a las campañas, son los costos los que aumentan más que los montos en proporción.

Esta conclusión aporta a la noción de que no siempre es lo óptimo contactar a todos los clientes, ya que en algún momento los costos empiezan a crecer más que los montos de los productos. Si bien las cifras son despreciables, hay que tomar en cuenta que se está considerando sólo el costo directo de los toques de la campaña, sin considerar otras fuentes de costo que podrían acercar más el número al de los montos y poner en duda la decisión de si contactar a toda la base.

9.4 Comportamiento de LIFT

Como análisis final y tomando en cuenta los puntos anteriores, se procede a analizar los valores de LIFT en los deciles de clientes y sus variaciones. De este análisis se procede a definir puntos de corte de LIFT que clasifiquen a los clientes (por decil o porcentaje de población) en las categorías de influenciado, causa perdida/asegurado, o *sleeping dog*.

Se procede a construir la Tabla 17 usando los datos de todos los clientes de las campañas de diciembre y de enero por decil de LIFT, que ilustra el promedio, máximo, mínimo y la desviación estándar de LIFT de cada decil. Luego, para estudiar la variación de cada decil, se calcula para cada uno el valor del máximo LIFT sumado a la desviación estándar promedio (de valor 0,127) a través de todos los deciles, y la resta del mínimo LIFT de cada decil con esta desviación estándar promedio.

Deciles	Promedio	Máximo	Mínimo	Desv. estándar	Máx + Desv. promedio	Mín - Desv. promedio
1	1,832	6,299	0,985	0,753	6,426	0,858
2	0,659	0,985	0,443	0,155	1,112	0,315
3	0,333	0,443	0,246	0,056	0,570	0,119
4	0,187	0,246	0,137	0,031	0,373	0,010
5	0,098	0,137	0,062	0,022	0,264	-0,065
6	0,032	0,062	0,004	0,017	0,189	-0,123
7	-0,020	0,004	-0,043	0,014	0,131	-0,171
8	-0,067	-0,043	-0,092	0,014	0,084	-0,220
9	-0,127	-0,108	-0,174	0,023	0,019	-0,301
10	-0,349	-0,174	-2,350	0,187	-0,047	-2,477

Tabla 17: Análisis de distribución de valores de LIFT en los clientes de campañas de enero y diciembre. Los valores de LIFT se han amplificado por 100 para facilitar la lectura y el análisis.

Al analizar los datos, se puede notar que el único decil que contiene valores de LIFT positivos y negativos es el séptimo, pero tomando en cuenta las dos últimas columnas, (que se han coloreado según si los valores son mayores a cero o no) se observa que hay valores mayores y menores a cero entre los deciles 5 y 9, si bien en el noveno el valor de Más + Desv. promedio es superior a cero por muy poco, de hecho por menos que la desviación estándar del decil mismo. Tomando esto en cuenta, se propone según la clasificación de Uplift Modelling definir a los clientes de los primeros 4 deciles como clientes influenciables, a los clientes de los deciles 5 a 8 como clientes no-influenciables (que comprenden clientes *sure-things* y clientes *lost-cause* que son indistinguibles para efectos de valor de LIFT) y a los clientes de los últimos deciles, 9 y 10, como clientes *sleeping dogs*.

Luego, si nos referimos a los resultados del análisis presentado en la sección 8.4 del documento, específicamente en la Tabla 10, se puede encontrar un respaldo para esta clasificación. Si se toman las tasas de respuesta según porcentaje de la población contactada del sistema de priorización por LIFT (tanto por sí solas como con el efecto *sleeping dogs* considerado) y se analiza el crecimiento de la TR con respecto al decil anterior, como se ve en la Tabla 18, se puede ver que es al pasar a los dos últimos deciles (del 80% al 90% y del 90% al 100% de la población, que corresponderían a los deciles 9 y 10 de la clasificación de LIFT) es cuando la tasa de respuesta incremental empieza a ser negativa. Por lo tanto, se apoya la decisión de establecer esos dos últimos deciles como los clientes perros dormidos.

Porcentaje	TR	Diferencia	TR+SD	Diferencia
10%	0,09%	-	0,21%	-
20%	0,41%	0,32%	0,54%	0,33%
30%	0,59%	0,18%	0,71%	0,17%
40%	0,7%	0,11%	0,82%	0,11%
50%	0,73%	0,03%	0,86%	0,04%
60%	0,79%	0,06%	0,92%	0,06%
70%	0,8%	0,01%	0,89%	-0,03%
80%	0,72%	-0,08%	0,78%	-0,11%
90%	0,55%	-0,17%	0,58%	-0,20%
100%	-	-	-	-

Tabla 18: Resultados de sistema de priorización por LIFT en campaña de enero, con el cálculo de la diferencia en TR al pasar de un escalón al siguiente.

Por otra parte, se puede observar que en los primeros cuatro deciles es donde se observa el crecimiento más grande en la TR al pasar de un escalón de porcentaje al siguiente, hasta llegar al paso del 40% al 50% en donde la tasa sólo aumenta un 0,03%, y no aumenta mucho más después. Estos datos reforzarían la idea de establecer esos deciles, del 5 hasta el 8, como los clientes no-influenciables, ya que agregarlos a la población contactada tiene poca influencia en la tasa de respuesta comparado con los clientes influenciados de los primeros 4 deciles.

CONCLUSIONES

A pesar de haber encontrado dificultades en el proceso, el proyecto pudo seguir su curso planeado en sus actividades más importantes, como el acceso a los datos y los experimentos de campo. Por otro lado, una preocupación previa a aplicar estos experimentos era que la tasa de respuesta fuera tan baja que el poder estadístico de la muestra fuera muy bajo y no permitiera concluir. El hecho es que en ambos experimentos se logró superar con creces la tasa de respuesta mínima necesaria, lo que asegura resultados estadísticamente significativos. Una razón para estas preocupaciones es que febrero y marzo son meses particulares, que son transaccionalmente distintos a otros. Otra preocupación tiene que ver con la contingencia actual, en la que se suman la crisis social y política de los últimos meses del 2019 y la nueva crisis social y de salud del 2020 por la pandemia mundial causada por el virus Covid-19. A pesar de estas situaciones, se logró alcanzar y superar la tasa de respuesta necesaria.

También de este hecho se puede concluir en primera instancia que los experimentos de campo pueden funcionar a pesar de situaciones extraordinarias, y que dado su potencial de aprendizaje se podrían utilizar en mayor medida para adquirir insights importantes y acercarse más al objetivo de conocer más a los clientes y poder entregar un mejor servicio.

En línea con lo mencionado, una propuesta de testeo experimental de la clasificación de LIFT sería llevar a cabo una campaña a un grupo de clientes escogidos según su LIFT (por ejemplo, el 10% de clientes con mejor LIFT) y a otro grupo separado de clientes escogidos aleatoriamente o bajo un sistema de elección definido por la empresa. Estos dos grupos, sumados a un grupo de control que no recibiría ninguna campaña, podrían recibir las campañas durante un mes y se podrían comparar los resultados entre ellos, y evaluar el impacto del modelo de LIFT de manera experimental, con la certeza que entrega este enfoque y la línea base que proporciona el grupo de control. También se convierte en un experimento fácil de replicar con variaciones simples pero importantes, como el porcentaje de población a publicitar, o los canales a utilizar.

El punto de los distintos canales o combinaciones de ellos para hacer llegar las campañas es especialmente interesante de seguir estudiando, ya que, si bien se ha demostrado con los experimentos que en promedio la tasa de respuesta es mejor publicitando por todos los canales que sólo por call-center, y que ambos tratamientos superan al tratamiento de email-SMS-carta, también se puede observar en las curva de Qini de estos tratamientos que el mejor resultado esperado no siempre es publicitar a través de todos los canales a

todos los clientes. En el Gráfico 7 se puede apreciar que lo óptimo al tener un grupo de clientes al que se va a llamar sería dirigir además publicidad multi-canal sólo a un 40% o 50% de los clientes más receptivos, con el mayor valor de LIFT, mientras que hay más clientes menos receptivos que son más susceptibles a molestarse al agregar más publicidad además de la llamada. Por otro lado, en el Gráfico 8 se puede observar la conclusión análoga, en el cual se indica que al haber definido un grupo de clientes que van a ser contactados a través de e-mail, SMS o carta, lo óptimo sería además publicitar a través de llamada a un 50% de los clientes más receptivos.

Luego, tomando el análisis previo en consideración y para que la empresa pueda utilizar el sistema de LIFT de una manera simple y útil, se decide establecer una segmentación en base a LIFT que sigue los siguientes principios:

- Los clientes con un LIFT mayor a 0,001 se consideran como *influenciables*.
- Los clientes con un LIFT menor a 0,001 y mayor a -0,001 se consideran como *no-influenciables*.
- Los clientes con un LIFT menor a -0,001 se consideran como *perros dormidos*.

Con estos puntos de corte, se toma a grandes rasgos la recomendación de la sección 9.4, es decir, los clientes *influenciables* contienen a los clientes de los primeros 4 deciles, los clientes *no-influenciables* contienen a los clientes de los siguientes 4 deciles, y los clientes *perros dormidos* contienen a los clientes de los últimos 2 deciles, con algunos clientes en los deciles de corte que entran en una categoría u otra depende de su cercanía con el punto de corte.

De esta manera, y con los datos utilizados para las campañas estudiadas, ya se cuenta con una clasificación para 848.397 clientes, de los cuales 366.329 están clasificados como *influenciables*, 320.265 como *no-influenciables* y 161.803 como *perros dormidos*. Además, de querer estimar el LIFT de más clientes, el modelo sólo necesita las covariables explicitadas en la sección 8.1 para entrega el resultado del LIFT de clientes nuevos.

Como se ha analizado en la sección 8.4, el sistema recomendado para priorizar a los clientes a los que se les va a dirigir campañas de marketing sería publicitar primero a los clientes *influenciables*, luego si se desea aumentar la población contactada, seguir con los clientes *no-influenciables*, priorizando dentro de las segmentaciones siempre por el valor de LIFT. Finalmente, es recomendable no hacer llegar campañas a los clientes catalogados como *perros dormidos*.

Por otra parte, si bien se ha realizado un análisis superficial del comportamiento de pago en la sección 9.1 del documento, es un aspecto interesante de estudiar en más profundidad, ya que es clave para el negocio de la empresa, sobre todo si se plantea relacionarlo con una metodología de uplift dirigido a campañas de pago de deuda o repactación de deuda, que están directamente relacionadas al comportamiento de pago.

Se podría replicar la metodología desarrollada en este proyecto experimentando con este tipo de campañas que actualmente la empresa realiza, y medir el pago de la deuda posterior de los clientes como la medida de respuesta, entrando quizás en más profundidad que una variable binaria de si un cliente paga o no paga. Se podrían analizar montos de pago, regularidad de pago, relación de pago del cupo normal de la tarjeta contra clientes que toman avances o súper avances, y más variables. Es interesante la noción de descubrir qué clientes son los más o menos receptivos a campañas de este tipo, y crear una segmentación más específica, en el sentido de que quizás a la empresa no le interesa sólo encontrar a los clientes que pagan su deuda cuando son contactados, sino a los que pagan en montos o porcentajes específicos de su deuda al ser o no contactados con una campaña con ese fin. En esta línea se podría estimar un LIFT de primer orden que mida la receptividad de los clientes a pagar deuda cuando son contactados, y luego aplicar un segundo filtro que segmente a los clientes en razón de qué tan interesantes son para la empresa desde el punto de vista de pago de deuda, y de esta manera priorizar a los clientes para dirigir las campañas de marketing de manera óptima.

Siguiendo esta línea, se podría combinar el conocimiento del cliente extraído de este proyecto, que busca maximizar el impacto de una campaña de productos financieros, con el conocimiento del cliente sobre su comportamiento y probabilidades de pago de deudas tanto de su cupo de tarjeta como de estos productos financieros.

Finalmente, el objetivo del proyecto se pudo cumplir. Se ha diseñado la metodología a través del modelo de predicción de LIFT, y su aplicación a una segmentación en tres grupos permite dirigir las campañas de marketing de la empresa con un conocimiento acerca de la receptividad de cada cliente al momento de ofrecer un producto financiero.

BIBLIOGRAFÍA

- [1]. Stiller, Ben (2018). Insight drive organization. Putting data-driven insights to work everywhere, everyday. Deloitte.com [en línea]. <https://www2.deloitte.com/us/en/pages/deloitte-analytics/solutions/insight-driven-organization.html>
- [2]. Ascarza, Eva (2018). Retention Futility: Targeting High-Risk Customers Might Be Ineffective. Columbia Business School Research Paper No. 16-28. Disponible en SSRN: <https://ssrn.com/abstract=2759170>
<http://dx.doi.org/10.2139/ssrn.2759170>
- [3]. Radcliffe, N.J., & Surry, P.D. (2012). Real-World Uplift Modelling with Significance-Based Uplift Trees. En Stochastic Solutions, sección Papers: <http://www.stochasticsolutions.com/pdf/sig-based-up-trees.pdf>
- [4]. Ascarza, Eva (2018), op. cit.
- [5]. Rubin, Donald (2005). Causal Inference Using Potential Outcomes. Journal of The American Statistical Association 100, 322-331. DOI: 10.1198/016214504000001880.
- [6]. Guelman, Leo, Guillén, Montserrat & Perez-Marin, Ana (2015). Uplift Random Forests. Cybernetics and Systems 46. DOI: 10.1080/01969722.2015.1012892.
- [7]. Rzepakowski, Piotr & Jaroszewicz, Szymon (2012). Decision trees for uplift modeling with single and multiple treatments. Knowledge and Information Systems - KAIS 32. DOI: 10.1007/s10115-011-0434-0.
- [8]. Breiman, L. (1996). Bagging Predictors. Machine Learning 24, 123-140 DOI:10.1023/A:101805431435.
- [9]. Guelman, Leo & Guillen, Montserrat & Perez-Marin, Ana (2015), op. cit.
- [10]. Guelman, Leo & Guillen, Montserrat & Perez-Marin, Ana (2015), op. cit.
- [11]. Guelman, Leo (2013). uplift Documentation. Disponible en <https://rdr.io/cran/uplift/man/upliftRF.html>
- [12]. Karlsson, Henrik (2019). Uplift Modeling: Identifying Optimal Treatment Group Allocation and Whom to Contact to Maximize Return on Investment (Dissertation). Disponible en <http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-157962>

ANEXOS

Anexo 1: Estadísticas de LIFT por decil de uplift bins.

Decil	Promedio	Máximo	Mínimo	Mediana	Desviación estándar
1	0,0285	0,1123	0,0123	0,0186	0,0223
2	0,0093	0,0123	0,0070	0,0091	0,0015
3	0,0056	0,0070	0,0044	0,0055	0,0007
4	0,0037	0,0044	0,0030	0,0036	0,0004
5	0,0025	0,0030	0,0020	0,0025	0,0003
6	0,0017	0,0020	0,0014	0,0017	0,0002
7	0,0011	0,0014	0,0009	0,0011	0,0001
8	0,0007	0,0009	0,0004	0,0007	0,0001
9	0,0001	0,0004	-0,0004	0,0001	0,0002
10	-0,0025	-0,0004	-0,0211	-0,0015	0,0025

Anexo 2: Resultados modelo predicción de pago de deuda

	<i>Variable dependiente:</i>
	PAGO
AVANCISTA1	1.130*** (0.161)
MONTO.Sin.PREPAGO	0.00000*** (0.000)
CCSAE.COLOR.NUEVOM	-0.031 (0.052)
CCSAE.COLOR.NUEVONI	-0.285*** (0.066)
CCSAE.COLOR.NUEVOR	-0.268*** (0.060)
CCSAE.COLOR.NUEVOV	-0.202*** (0.041)
scoreCCSAE.MR	0.027*** (0.0003)
CUPO_TOTAL	-0.00000*** (0.000)
DIAS.MORA	-0.033*** (0.006)
HABILITADO.COMPRA1	2.222*** (0.201)
HABILITADO.REPA1	-1.870*** (0.130)
TCA1	0.809*** (0.033)
TCA2	1.676*** (0.180)
DISP_AVANCE_TOTAL	0.00000*** (0.000)
Constante	-8.776*** (0.207)
2	
Observaciones	66,136
Log Likelihood	-34,666.110
Akaike Inf. Crit.	69,362.230
<i>Nota:</i>	*p<0.1; **p<0.05; ***p<0.01

Anexo 3: Matriz de confusión datos de entrenamiento

	No paga	Paga
No paga	15.478	6.357
Paga	10.420	33.881

Anexo 4: Matriz de confusión datos de testeo

	No paga	Paga
No paga	3.851	1.571
Paga	2.710	8.403
