

UNIVERSIDAD ADOLFO IBÁÑEZ

Segmentación de leads a partir del valor potencial de los clientes para incrementar las ventas.

4 DE JULIO, 2023

Autor: Sebastián Higuera
Carrera: Ingeniería Civil Industrial
Profesor: Francisco Duque
Supervisor: Diego Fernandez

Índice:

Índice:	2
1. Resumen ejecutivo:	3
2. Abstract:	4
3. Introducción:	5
a. Contexto de la empresa	5
b. Contexto del problema	5
4. Objetivos:	10
a. Objetivo general	10
b. Objetivos específicos y medidas de desempeño	11
5. Estado del arte:	13
a. Marco teórico	13
b. Alternativas de solución	14
6. Solución:	18
a. Metodologías	18
b. Solución alcanzada	18
7. Resultados:	22
a. Resultados	22
b. Impacto	23
8. Conclusión:	26
9. Discusión:	27
10. Referencias:	29
11. Anexos:	30
Anexo A	30
Anexo B	31
Anexo C	33
Anexo D	36
Anexo E	38

1. Resumen ejecutivo:

Simplee es una Start-up con 5 años en el rubro de las corredoras de seguros que se especializa en pólizas de seguros para PYMEs con un foco innovador y ágil para otorgar un mejor servicio y acompañar a aquellas personas que han decidido emprender. Simplee cuenta con su sede principal en Santiago de Chile y una segunda sede en México.

Dentro de esta empresa, se llevó a cabo un análisis del funcionamiento de los procesos, costos, ventas y marketing ambientado al área de Business Analytics, área donde se trabajó y llevó a cabo el proyecto. A partir de este análisis, se encontró que existe una diferencia de 5795 ventas versus 13793 leads que iniciaron el proceso de cotización, lo que indica que la brecha de gente que inicia el proceso de cotización y que finalmente resulta en una venta es de un 42%. Esta cifra indica un claro problema, ya que cada lead tiene un costo de adquisición de 3 USD.

Con el objetivo de reducir esta brecha entre la cantidad de ventas y la cantidad de gente que comienza el proceso de cotización, para de esta manera aumentar la ventas y reducir las pérdidas por adquisición de leads que no producen ganancia, se realiza un modelo de machine learning que entrega el valor potencial de los clientes, reduciendo la incertidumbre, permitiendo segmentar a los clientes y permitiendo realizar campañas de Cross-Selling a partir de los datos obtenidos del proyecto, dando una oportunidad de vender más productos basados en el giro de la empresa y el o los productos comprados anteriormente.

Se logró realizar un sistema que entregue el valor potencial de nuevos clientes basados en las variables claves (giro, compañía, comuna, producto, grupo de producto y estado) y que permite segmentar a los clientes para entregar un mejor y más detallado servicio a estos. Además, se logró realizar una campaña de Cross-Selling con éxito, permitiendo entregar un mejor servicio a los clientes y aumentar las ventas sin incurrir en nuevos gastos por adquisición de nuevos clientes.

2. Abstract:

Simplee is a 5-year-old start-up in the insurance brokerage industry that specializes in insurance policies for SMEs (Small and Medium Enterprises) with an innovative and agile focus to provide better service and support for those who have decided to embark on entrepreneurship. Simplee has its main headquarters in Santiago, Chile, and a second branch in Mexico.

Within this company, an analysis of the functioning of processes, costs, sales, and marketing was carried out, focused on the Business Analytics area, where the project was developed and executed. From this analysis, it was found that there is a difference of 5795 sales compared to 13793 leads that initiated the quotation process, indicating a 42% gap between the number of people who start the quotation process and those who ultimately result in a sale. This figure indicates a clear problem, as each lead has an acquisition cost of \$3 USD.

With the objective of reducing this gap between the number of sales and the number of people who initiate the quotation process, in order to increase sales and reduce losses from leads that do not generate profit, a machine learning model was developed to provide the potential value of customers, reducing uncertainty and allowing for customer segmentation and cross-selling campaigns based on the data obtained from the project. This provides an opportunity to sell more products based on the business sector and the previously purchased product(s).

A system was successfully developed to provide the potential value of new customers based on key variables (sector, company, locality, product, product group, and status), allowing for customer segmentation to provide a better and more detailed service. Additionally, a successful cross-selling campaign was conducted, enabling the delivery of better service to customers and increasing sales without incurring new expenses for acquiring new customers.

3. Introducción:

a. Contexto de la empresa.

Con más de 5 años en el mercado, Simplee es una corredora de seguros enfocada en emprendedores y Pymes de Chile y México. El modelo de negocio de Simplee es ofrecer una vasta variedad de pólizas de seguro proveniente de varias empresas aseguradoras asociadas que otorgan una comisión por venta realizada a través de la página web. Con la experiencia que tiene la empresa en el área de emprendimiento y siendo una Start-up, Simplee es capaz de ponerse en el lugar de sus clientes, entender sus problemáticas y solucionarlas. En la página web se puede elegir entre la variedad de seguros que ofrece la empresa, cotizar y ver la póliza de seguro que más le convenga a la empresa. (Más información sobre los productos que ofrece la empresa en el Anexo A).

La misión de Simplee es hacer del mundo de los seguros un mundo más asequible y ágil para las empresas y pymes. La manera en que piensa hacer esto es entregando propuestas de seguros que se acoplen a las necesidades y problemáticas de las empresas y pymes a partir de la experiencia y con un punto de vista de emprendimiento e innovación.

Al tratarse de una Start-up es evidente que todavía sigue en crecimiento y con constantes cambios, pero al mismo tiempo cuenta con 5 años de experiencia e información que ha ido recopilando durante este tiempo. Por otro lado, el foco de la empresa se encuentra en la innovación y entregar un servicio de calidad, por lo que realizar un proyecto innovador y pensado para mejorar la experiencia del cliente resulta conveniente en una empresa como Simplee.

b. Contexto del problema.

Al ser una Start-up, el crecimiento de Simplee específicamente en los últimos 3 años ha sido exponencial hasta el punto en que el número de ventas se estabilizó. La empresa se encontraba en pleno desarrollo de nuevas tecnologías y sistemas cuando debió cerrar el área de innovación y desarrollo debido al cierre del Silicon Valley Bank, que se llevó consigo a varios de los inversionistas de la empresa. Esto significó un fuerte golpe para la empresa, por lo que como se trata de una corredora de seguros que gana dinero por comisión de venta, se ven obligados a vender una mayor cantidad de seguros.

Esto lleva al principal síntoma que demuestra que existe un grave problema respecto a la brecha que existe entre los leads¹ que inician el proceso de cotización de pólizas de seguros y los que finalmente compran una póliza. La brecha que existe es de un 42% de

¹ Lead: Potencial cliente, persona que comienza proceso de cotización y puede ser un contacto de venta.

clientes en comparación con los leads que iniciaron el proceso.

Algunos de los Síntomas que explican el número de ventas de la empresa son:

- Alrededor de 25 empresas han comprado pólizas al menos una vez y sólo 8 de estas 25 empresas son consideradas clientes frecuentes de Simplee.
- Existen más de 13000 leads en las bases de datos que comienzan el proceso de cotización.
- El número de ventas realizadas son de 5795 pólizas de seguros.
- El alcance de las campañas de correos ha sido bajo, con un 38.8% de emails abiertos y 1.9% de los correos abiertos obtuvo al menos una interacción.

Estos son claros factores de que no se está aprovechando al máximo el potencial de ventas de la empresa y que existe una enorme brecha entre leads y clientes. Para entender mejor el proceso por el que pasa el cliente, se analiza en detalle este proceso en el Anexo B.

Para analizar las causas del problema se tuvo en cuenta algunos de los problemas que existen actualmente en la empresa en todo lo que esté relacionado con el manejo de datos y de información de los clientes debido a que el proyecto que se llevará a cabo corresponde al área de Business Analytics.

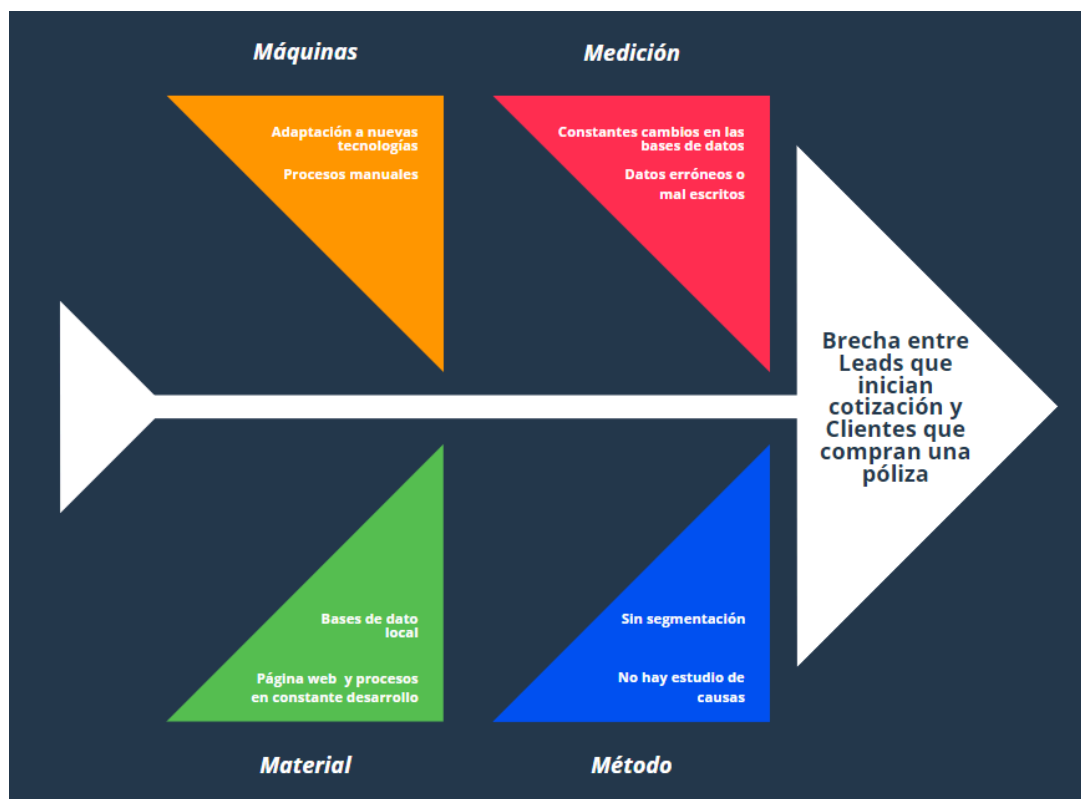


Figura 2: “Diagrama de Ishikawa para descubrir las posibles causas del problema”

Fuente figura 2: “Creación propia a partir de datos entregados por el supervisor Diego Fernandez, Augusto Carrillo del área de TI y Domingo Mederos encargado de

comunicaciones y marketing”

Utilizando la metodología de Ishikawa se encuentran algunas de las posibles causas, separadas en diferentes categorías, de por qué existe una brecha entre la cantidad de Leads que inician el proceso de cotización y los Clientes que finalmente compran una póliza de seguros.

En cuanto a lo relacionado con la maquinaria y material, algunos de los factores que más influyen en generar esta brecha son:

- Adaptación a nuevas tecnologías, al ser una empresa que se encuentra en constante crecimiento e innovación, se utilizan nuevas tecnologías constantemente para realizar un mejor trabajo y entregar un mejor servicio, por lo que tanto las personas del equipo que se encargan de todo lo que sucede por detrás en la empresa como las personas encargadas de realizar las ventas (KAM) deben ir adaptándose a estos cambios.
- Procesos manuales, relacionado a lo anterior, actualmente varios de los procesos que se realizan en la empresa son manuales. Los procesos que se realizan de manera manual hasta el momento son: transcripción a bases de datos de las KAM, atención a los clientes, contacto con los clientes, renovación de pólizas, servicios post venta. Por otro lado, las campañas de correos se encuentran en proceso de automatización.
- Bases de datos local, como se mencionó anteriormente, las bases de datos actualmente se realizan de manera manual y local, es decir, durante el día cada KAM escribe todo en un excel local y luego se traspasa todo a una base de datos online.
- Página web y procesos en constante desarrollo, esto provoca que la cantidad de información que se tiene de años anteriores y la que se tiene actualmente sea totalmente distinta y que estandarizar la información resulte más complejo, además de ser el cambio más notorio para los Leads y su comportamiento.

Respecto a la medición:

- Constantes cambios en las bases de datos, esto provoca que haya vacíos e irregularidades, por lo que dificulta utilizar esta información recopilada para incrementar las ventas.
- Datos erróneos o mal escritos, al no estar estandarizadas las variables y hacer los trasposos de bases de datos locales a bases de datos globales suele suceder que ocurren errores o se escriben de distintas maneras los datos. Algunos ejemplos de esto y que demuestran la gravedad de esto para realizar una predicción adecuada utilizando la base de datos de ventas es, en primer lugar, en la variable Comuna

existen 426 variables únicas siendo que existen solo 346 comunas en todo Chile, además de contar con 607 datos vacíos o nulos; por otro lado, la variable Giro cuenta con 751 valores únicos y 919 datos vacíos o nulos.

Respecto al método:

- Sin segmentación, no existe actualmente un método que diferencie un cliente/lead de otro, por lo cual se ofrecen los mismos productos y nivel de atención a todos por igual.
- No hay estudio de causas, no hay conciencia de que existe este problema y por lo tanto no se han tomado las acciones necesarias para evitar que esta brecha siga creciendo.

El costo de adquisición de clientes (CAC) promedio entre los meses de Julio 2022 hasta Marzo 2023 son:

- CLP= 3 USD Costo de campañas (Google Ads y RRSS) por leads obtenidos.
- Performance CAC= 9 USD Costo de campañas (Google Ads y RRSS) por pólizas vendidas.
- Blended CAC= 26 USD Costo de campañas (Google Ads, RRSS y Gasto comisión) por pólizas vendidas.

Se evaluó en estos 3 CAC distintos debido a que existen varias maneras distintas de evaluar el CAC según los tipos de gastos en que se incurran y que factores se quieran tomar en cuenta para este cálculo.

Por otro lado, al analizar la base de datos de ventas históricas hasta marzo de 2023, se obtiene que la moda de la comisión final fue de 1.6983 UF de un total de 5700 pólizas vendidas. Por lo tanto, se puede decir que por cada póliza que no se vende la empresa deja de recibir un estimado de 1.6983 UF, cifra no menor ya que cada Lead que no compra tiene un costo ponderado de 3 USD si no genera una venta. Siendo la brecha entre Leads y Clientes de 8000 personas que iniciaron el proceso de cotización pero no concluyeron la compra, se puede decir que el costo del problema es de $8000 \times 3 \text{ USD} = 24000 \text{ USD}$.

Para cuantificar el alcance a posibles Clientes y nivel de interés de los Leads se utilizarán los datos obtenidos de las campañas de correos que se han realizado, representado en los siguientes indicadores:

Los valores antes de aplicar el proyecto de estos KPI, evaluado en mayo de 2023 son:

% de alcance campaña de correos = 38.8%

Este indicador muestra que no se está pudiendo contactar a más de la mitad de los leads, por lo tanto resulta imposible ofrecer nuevos productos o renovación de productos a estos.

% de interacción campaña de correos = 1.9%

Este indicador muestra que del 38.8% que se contactó, sólo un 1.9% de ellos hizo click en algún link o resultó en una venta gracias a la campaña de correos.

Cantidad de ventas campaña Cross Selling = 2

Este indicador muestra los resultados de la última campaña de Cross-Selling que se realizó de manera general y el bajo rendimiento que obtuvo.

Para cuantificar la cantidad de leads que se han transformado en clientes se comparará la cantidad de datos en la base de datos de leads con la cantidad que se encuentran en la base de datos de clientes, lo cual entrega un total de 5795 clientes versus 13793 leads, es decir, sólo un 42% de los leads que inician el proceso de cotización de una póliza termina comprando una póliza.

4. Objetivos:

a. Objetivo general.

Reducir la brecha que separa la cantidad de leads que finalizan el proceso de cotización hasta comprar un producto, aumentando 1.5 puntos porcentuales la proporción de clientes/leads para final de Junio.

En un total de 49 meses se realizaron un total de 5795 ventas, es decir, se realizaron alrededor de 118 ventas por mes. Calculando el promedio de ventas desde el año 2021 que es el año en que la cantidad de ventas se estabilizó, se encontró que en 2021 se vendió en promedio 108 pólizas, en 203 en 2022 y 175 en 2023 hasta el mes de abril, por lo que en promedio se vende 162 pólizas por mes desde que las ventas se estabilizaron. De esta manera, si se aumenta 1.5 puntos porcentuales la proporción de clientes/leads en un mes, considerando como base 13793 leads actuales, se estaría consiguiendo un total de 205 nuevas ventas en un mes, 26.5% más que el promedio por mes que se tiene actualmente. La idea de este proyecto es conseguir más clientes a partir de los leads que se tienen, por lo que la brecha existente debería disminuir, aumentando la cantidad de clientes y manteniendo la cantidad de leads existentes.

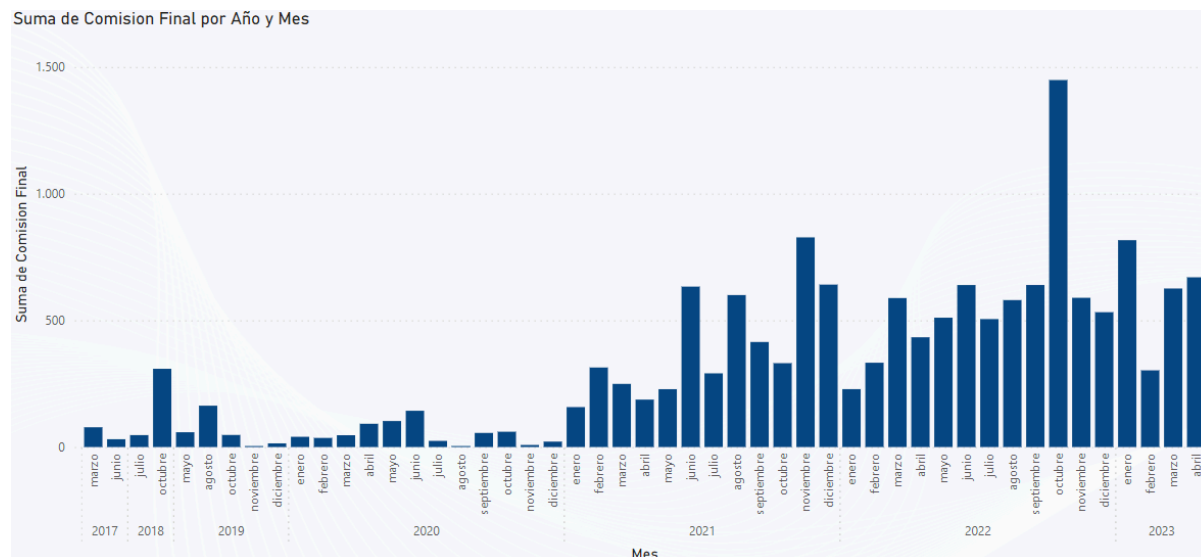


Figura 3: “Gráfico que muestra la suma de Comisión final obtenida por mes desde marzo de 2017 hasta abril de 2023”

Fuente figura 3: “Creación propia a partir de la base de datos de ventas de Simplee con Power BI”

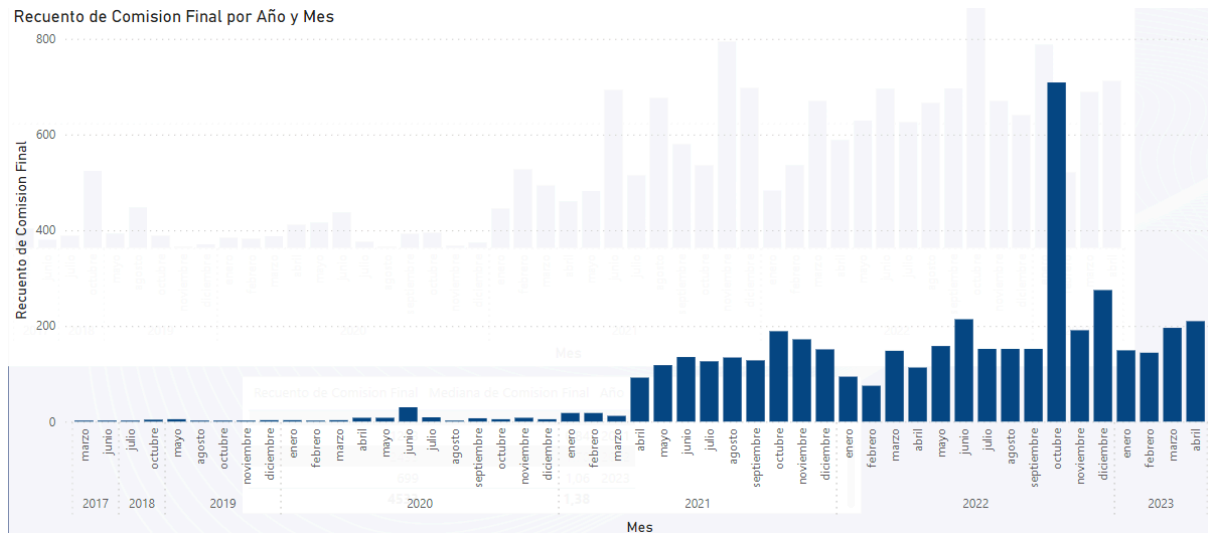


Figura 4: “Gráfico que muestra la cantidad de ventas obtenidas por mes desde marzo de 2017 hasta abril de 2023”

Fuente figura 4: “Creación propia a partir de la base de datos de ventas de Simplee con Power BI”

Para medir si el objetivo general se cumplió con éxito se utilizará el siguiente KPI:

$$(F1) \% \text{ de clientes por lead} = \frac{\text{Cantidad de Clientes}}{\text{Cantidad de Leads}} * 100$$

Como el valor antes de aplicar el proyecto, medido en el mes de mayo de 2023 con datos históricos desde septiembre de 2017 hasta abril de 2023, es de un 42%, el valor óptimo para decir que el proyecto tuvo éxito es de un 43.5% en este KPI.

b. Objetivos específicos y medidas de desempeño.

Para cumplir con el objetivo general se debe alcanzar primero los siguientes objetivos específicos:

- Aumentar el alcance de la campaña de correos a un 50%.
- Alcanzar un 4% de interacción con las campañas de correos.
- Conseguir al menos 15 ventas² a leads que se hayan generado por la campaña de Cross-Selling.

Para medir si los objetivos específicos se cumplieron con éxito se utilizará los siguientes KPI:

² Se pide al menos 15 ventas para cubrir los gastos de mano de obra considerando un ingreso de 1.6983UF por venta realizada promedio.

$$(F2) \% \text{ de alcance campaña de correos} = \frac{\text{Cantidad de Emails abiertos}}{\text{Cantidad de Emails enviados}} * 100$$

$$(F3) \% \text{ de interacción campaña de correos} = \frac{\text{Cantidad de interacciones}}{\text{Cantidad de Emails abiertos}} * 100$$

$$(F4) \text{ Cantidad de ventas campaña Cross Selling} = \Sigma(\text{ventas Cross Selling})$$

5. Estado del arte:

a. Marco teórico.

Actualmente no existe una noción de lo grande que es esta brecha que existe entre los leads que inician el proceso de cotización y los que finalmente compran un producto, pero de todas maneras la empresa ha realizado distintas acciones para aumentar las ventas. Estas acciones son:

- Plan de referidos: Se ofrece una parte de la comisión de la venta a la empresa que referido a Simplee y haya finalizado en la venta de un producto.
- Campañas de correo automatizadas: Se generó una campaña de correos para ofrecer los productos de Simplee de manera automática y agregando los datos de las KAM para ser contactadas en caso de interés.
- Alianza con aseguradoras: Buscar mejores acuerdos y ofertas con las aseguradoras asociadas con Simplee para ofrecer mejores productos y ofertas a los clientes.

De acuerdo a la investigación realizada anteriormente (investigación presente en el Anexo C) del estado del arte para la solución de problemas de fidelización del cliente o conversión de leads, se formularon las siguientes alternativas de soluciones adaptadas a esta empresa y considerando ideas existentes mezcladas con la investigación mencionada anteriormente:

- **Identificación de leads según su potencial de convertirse en clientes.** Se realiza un modelo de machine learning con la data histórica de la empresa para hacer un sistema que permita predecir qué tan probable es que cada lead/cliente realice una compra de uno o varios productos de la empresa. Esto permite poner un especial énfasis en los potenciales clientes y ofrecer productos que se acoplen mejor a las necesidades de estos. Además permite encontrar patrones sobre los productos y giro de las empresas que tienen una mayor probabilidad de convertirse en clientes.
- **Agregando valor al producto ofrecido.** Manteniendo un conocimiento actualizado de los productos y servicios de seguros que se encuentran disponibles en el mercado, de esta manera se podrá ofrecer opciones de seguros más adecuadas y competitivas a los clientes, y también ayudará a identificar oportunidades para upselling o cross-selling, lo que podría aumentar los ingresos y comisiones de la empresa.
- **Campañas de marketing automatizadas.** Esto permitirá llegar a un mayor número de personas para mostrar los servicios y productos que ofrece la empresa y de esta manera aumentar las posibilidades de generar una venta.
- **Personalización de mensajes.** Al hacer llegar mensajes más personalizados a los

leads, se está mostrando un mayor interés y mejor servicio de parte de la empresa hacia los posibles compradores. Esto generará una mayor satisfacción respecto al servicio e imagen de la empresa y entregará una mejor experiencia de consumo, incentivando a volver a comprar en un futuro con la misma empresa.

- **Seguimiento activo de clientes potenciales por canal de preferencia.** Al igual que el anterior, este plan de acción busca mejorar la experiencia de consumo al mostrar interés en el lead y generar un contacto más eficaz al contactarse por la fuente de contacto de preferencia y enfocándose en los potenciales compradores.
- **Simplificando el flujo de interacciones.** Se hacen cambios en la página para hacerla más atractiva y amigable visualmente para el consumidor, se crea una opción de crear un usuario para guardar los datos de la empresa/persona natural y así evitar tener que llenar todos los datos cada vez que se quiera cotizar un producto, reduciendo el esfuerzo que debe hacer el consumidor y mejorando la experiencia de compra y servicio. Para esto no se necesita realizar gastos extra ya que se puede realizar en conjunto con el departamento de TI, pero significaría una reestructuración completa de la página y de las bases de datos.

b. Alternativas de solución.

Dentro de las alternativas mencionadas anteriormente en el marco teórico, se seleccionaron las 3 alternativas que cumplen con un desafío ingenieril, que permita tomar decisiones que ayuden a lograr el objetivo general y que permitan cumplir con los objetivos específicos. Luego, para medir cuál de estas alternativas se ajusta mejor al problema que se quiere solucionar, se realizó una tabla para analizar esto considerando razones de costo (cuán costoso sería realizar este proyecto y que tan dispuesta estaría la empresa de pagar esto), tiempo (cuánto tiempo tardará en realizarse este proyecto y si es posible realizarse en un tiempo establecido menor o igual a la duración de la pasantía), impacto (que tan grande sería el impacto en la empresa y que tanto influye en los objetivos establecidos) y técnico (que tan factible es realizar este proyecto con los conocimientos adquiridos y las herramientas disponibles).

Adecuación y Utilidad (0: No se adecúa o no es de utilidad al problema; 5: Se adecúa a la perfección y es de utilidad para resolver el problema planteado)					
Soluciones	Costo	Tiempo	Impacto	Técnico	Total
Agregación de valor al producto ofrecido.	5	3	4	5	17
Identificación de leads por valor potencial.	5	4	5	4	18
Simplificación de flujo de interacciones.	4	4	5	3	16

Tabla 1: “Tabla que mide que tan bien se adecúa y resuelve el problema planteado cada solución propuesta”

Fuente tabla 1: “Creación propia a partir de diálogos con el equipo de trabajo de Simplee”

Las 3 alternativas de solución se encuentran explicadas más a detalle en el Anexo D.

La solución escogida para realizar como proyecto para la empresa es la de “Identificación de leads según su potencial de convertirse en clientes”. Para seleccionar esta solución se tuvo en consideración los siguientes criterios:

1. Viabilidad técnica: Se considera viable en cuanto a lo técnico, ya que es posible implementar esta solución considerando las bases de datos de los últimos 5 años que se tienen, los programas para visualización y manejo de datos con los que cuenta la empresa, una suscripción en Customer IO para aplicar esta segmentación en futuras campañas de marketing y no cuenta con restricciones tecnológicas que impidan su funcionamiento.
2. Factibilidad operativa: Se considera que la solución seleccionada es factible operativamente, ya que puede implementarse en los procesos y operaciones que la empresa realiza actualmente sin afectar o complicar los flujos de trabajo actuales. El resultado se implementará como una segmentación en el sistema de Customer IO y además se utilizará para analizar patrones y comportamiento de potenciales clientes,

por lo que los demás trabajadores no verán su trabajo influenciado, y por lo tanto, su eficiencia y productividad tampoco se verán afectadas.

3. Impacto en los stakeholders: Se considera que la solución tendrá un impacto positivo en los stakeholders debido a que ayudará a segmentar los clientes y ofrecer un mejor servicio a los potenciales clientes, lo cual se alinea con los valores y objetivos de la empresa, sin afectar de ninguna manera a los empleados de Simplee.
4. Sostenibilidad y escalabilidad: Se considera que la solución es sostenible y escalable, ya que al ser una solución que viene de la data histórica de la empresa, no será necesario repetir el proceso hasta dentro al menos un año más, donde se deberá considerar la nueva data obtenida para realizar la segmentación de nuevo. Al estar todo programado, solo se deberá hacer correr el código una vez al año para revisar el resultado del modelo de Machine Learning y actualizar la plantilla que ayuda a segmentar los leads/clientes, todo lo demás se encontrará automatizado, por lo que el proyecto será repetible y con el feedback que se obtiene de los resultados, también es escalable.

Para evaluar los posibles riesgos de la solución escogida se analizará la probabilidad, impacto y prioridad de los riesgos encontrados al evaluar en primera vista los materiales con los que se cuentan y algunos escenarios hipotéticos.

Riesgo	Probabilidad	Impacto	Prioridad	Mitigación
Errores en las bases de datos	4	4	Alta	Realizar manipulación y limpieza de base de datos.
Pocos datos que entreguen valor	3	4	Alta	Realizar análisis de sensibilidad e identificar las variables que sean más relevantes para el proyecto.
Datos que no muestran un comportamiento normal	3	5	Alta	Realizar todos los análisis estadísticos necesarios y evaluar la posibilidad de tener que utilizar una base de datos distinta o múltiples bases de datos.
Leads/Clientes no interesados en otros productos	2	3	Media	Encontrar patrones de compras de Leads/Clientes similares y ofrecer otros productos adecuados (cross-selling) o mejoras en los productos que compran regularmente (Upselling).
Resultados imprecisos al segmentar	3	2	Media	Evaluar utilizar variables distintas y ajustes al modelo de machine learning.

Tabla 2: "Tabla de análisis de riesgo y mitigación"

Fuente tabla 2: "Creación propia a partir de comentarios de parte del supervisor Diego"

Fernandez, del equipo de Marketing y comentarios de estudiantes de Magíster de Data Science”

Como se puede observar de la Tabla 2, para este proyecto los mayores riesgos se ven presentes en los datos históricos con los que cuenta la empresa, ya que si estos están mal escritos, alterados o no muestran un comportamiento normal, el proyecto no podría llevarse a cabo o habría que buscar una alternativa distinta. Afortunadamente, si bien las bases de datos con las que cuenta la empresa necesitaron un arduo trabajo para utilizarse para predicción, se logró superar estas dificultades con éxito. Por otro lado, los dos riesgos de prioridad media no afectan el proyecto hasta una vez terminado el primer modelo, por lo que la manera de mitigar estos resulta más fácil y se pueden solucionar a partir de feedback.

6. Solución:

a. Metodologías.

Para desarrollar el proyecto se utilizará la metodología Waterfall, la cual consiste en un enfoque secuencial y lineal para el desarrollo de proyectos. Se considera utilizar esta metodología dado que se trata de un proyecto que se lleva a cabo en un periodo de tiempo no lo suficientemente largo para que ocurran cambios significativos para llevar a cabo con éxito el proyecto, los riesgos encontrados tampoco se ven afectados por factores temporales y el plan de acción es sólido y contempla realizar pruebas y ajustes luego de terminada la primera versión del proyecto. Los pasos a seguir utilizando esta metodología son los siguientes:

- **Plan:** Realizar un reconocimiento de la empresa, aprendiendo todo sobre los procesos y el funcionamiento del mercado de las corredoras de seguros; analizar los posibles problemas existentes en la empresa; definir los KPIs; y evaluar estado del arte para aumentar las ventas en corredoras de seguros.
- **Diseño:** Seleccionar una solución al problema planteado; seleccionar el modelo de Machine Learning que se adecúe mejor al tipo de datos trabajados para predecir.
- **Desarrollo:** Preparar bases de datos para aplicar el modelo de Machine Learning; ejecutar el modelo de Machine Learning elegido.
- **Test:** Aplicar el modelo resultante para evaluar el valor potencial de los clientes en datos históricos y evaluar el rendimiento y los resultados obtenidos.
- **Despliegue:** Aplicar el modelo resultante para acciones de venta y marketing.
- **Revisión:** Revisar los KPIs finales.
- **Feedback:** Utilizar la retroalimentación obtenida de la implementación del modelo para mejorar el modelo Machine Learning y realizar predicciones más adecuadas.

b. Solución alcanzada.

Para desarrollar la solución se siguieron los siguientes pasos (los detalles de todo el proceso se encuentra en el Anexo E):

1. **Análisis de la empresa y el mercado de corredoras de seguros:** Se investiga sobre el rubro, analizando cómo funcionan los procesos en la empresa y de qué manera genera valor y ganancias; y por otro lado, se investiga cómo trabaja la competencia y cómo se mueve la industria de las corredoras de seguros. Además se investiga qué es lo que motiva a las empresas a comprar seguros y qué factores y variables son las que más influyen en el precio de un seguro.

2. Obtención de bases de datos: Se analizaron todas las bases de datos con las que cuenta la empresa Simplee en búsqueda de la base de datos que tenga la menor cantidad de datos vacíos o errores y que tenga la mayor cantidad de variables que puedan tener un cierto grado de importancia para lograr predecir, basado en los resultados obtenidos en la investigación anterior.
3. Manipulación bases de datos: Se descarga la base de datos seleccionada ("Ventas_all"), se conecta a Power BI³ la base de datos para hacer una primera manipulación, haciendo algunos ajustes, reduciendo los errores y quitando algunas columnas que no serán de utilidad para el estudio. Luego se descarga la base de datos resultante de Power BI y se procede a trabajar en Python de Jupyter Lab⁴ para analizar paso a paso la base de datos. En este, se empieza a hacer los análisis estadísticos pertinentes (explicados en el Anexo E) para ver que la base de datos que se está trabajando efectivamente sirve para hacer una predicción, es decir, que la información muestre una distribución de comportamiento normal. Una vez demostrado esto, se procede a limpiar la base de datos, se eliminan los datos Outliers que no aporten en la predicción gracias al método de LOF⁵ y se hacen los ajustes necesarios para hacer una clusterización.

³ Se decide utilizar Power BI debido a su capacidad de visualización de datos y las facilidades para arreglar las bases de datos que entrega, en específico, se utilizó por la habilidad de arreglar errores de escritura en las bases de datos a tiempo real y para eliminar columnas innecesarias para el estudio sin afectar las bases originales.

⁴ Se decide utilizar Python de Jupyter Lab debido a las facilidades de análisis estadísticos paso a paso que otorga, además de contar con experiencia previa trabajando modelos de Machine Learning con este programa.

⁵ LOF: Local Outlier Factor, factor que compara la densidad de un punto de datos con sus vecinos para encontrar los valores atípicos y posteriormente eliminarlos de la muestra de datos para tener una predicción más precisa.

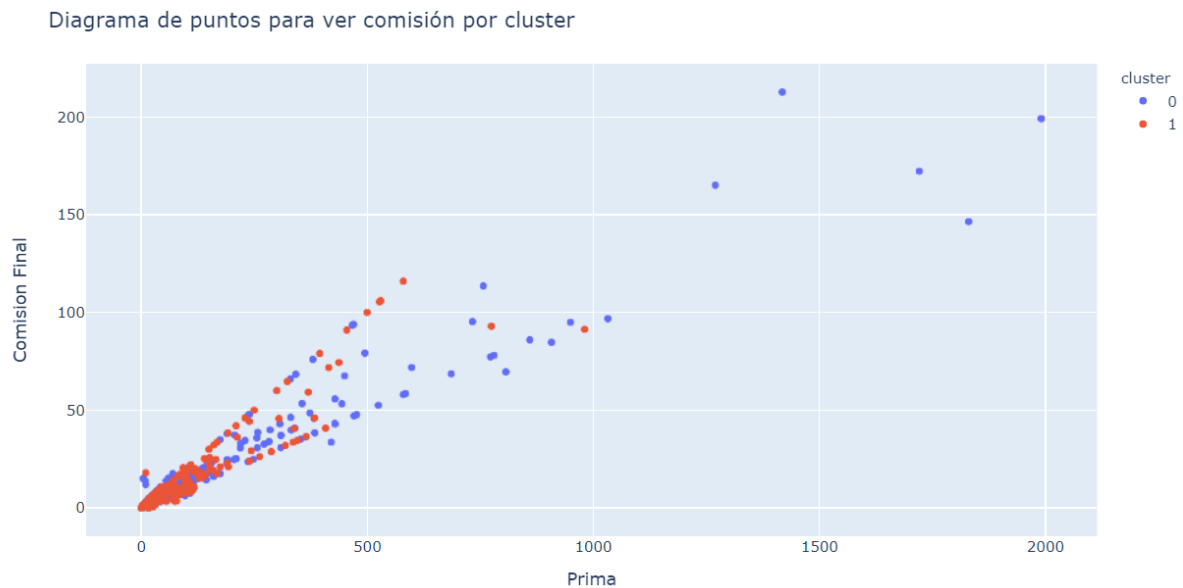


Figura 5: “Diagrama de puntos para visualizar la relación entre Comisión final y Prima segmentado por los clusters”

Fuente figura 5: “Creación propia a partir de la base de datos de ventas de Simplee con Python de Jupyter Lab”

4. Realización de modelo de Machine Learning: Se realiza un análisis de sesgo para revisar que el resultado de la predicción no resulte sesgado por ninguna de las variables principales. Luego se procede a seleccionar un modelo de Machine Learning, que en este caso al tratarse de variables categóricas se decide aplicar el modelo de árboles de decisiones. Se obtiene una imagen con el árbol de decisiones y se descarga un excel con la data clusterizada y otro con la data codificada.
5. Análisis árbol de decisiones y creación plantilla para calcular valor potencial de clientes: Se analiza el árbol de decisiones para crear unos filtros que luego se utilizan en Power BI con las bases de datos descargadas anteriormente. A partir de esto se hace una plantilla para excel que verifica si los leads cumplen con las condiciones que entrega el árbol de decisiones para calcular el valor potencial de estos.
6. Creación de segmentación: Se utiliza los resultados de esta plantilla para segmentar en 3 grupos los clientes, se separa en alto valor potencial (con puntaje 5 o 6), mediano valor potencial (con puntaje 4 o 3) y bajo valor potencial (con puntaje 2, 1 o 0).

Para implementar esto, se realizan los siguientes pasos:

1. Se importan las bases de datos segmentadas en Customer IO para crear segmentos para futuras campañas de correos.
2. Se crea un Diagrama de flujo para hacer campañas de Cross-Selling para cada segmento de valor potencial por Giro para hacer una campaña más precisa.
3. Se crea un mail para cada Giro que se quiera contactar ofreciendo los productos pertinentes.
4. Se envían los correos siguiendo el orden programado en Customer IO según los flujos creados para cada segmento de valor potencial.
5. Se obtiene feedback de los resultados obtenidos y se repite el proceso con los ajustes adecuados.

7. Resultados:

a. Resultados.

La implementación del proyecto del proyecto logró separar con éxito la base de datos en 3 segmentos de valor potencial (alto, mediano y bajo valor potencial), donde al ponerlos a prueba en una campaña de Cross-Selling, se obtuvo los siguientes resultados:

Campaña	Emails enviados	Emails abiertos	Interacciones	Cotizaciones/ Consultas
AVP ⁶	162	50,60%	3%	13
MVP ⁷	392	34,40%	1,30%	10
BVP ⁸	204	26,00%	0	2

Tabla 3: "Tabla comparativa de resultados campañas Cross-Selling"

Fuente tabla 3: "Creación propia a partir de datos entregados por Customer IO"

Si bien los resultados obtenidos de la campaña no son los esperados en cuanto a cantidad de ventas, se puede observar que la campaña que recibió un mayor interés por parte de los clientes es la de alto valor potencial (AVP) con un 50,6% de emails abiertos y un 3% de interacciones, seguido de la campaña de MVP y luego la campaña de BVP. Con propósito de estudiar la efectividad del sistema de puntuación de valor potencial se envió el mismo correo a los 3 segmentos de valor potencial para los 5 rubros que poseen un mayor registro de ventas en las bases de datos de Simplee (Negocios y finanzas; Servicios inmobiliarios, bodegas y estacionamientos; Gestión, asesoría y consultoría; Construcción y extracción; y Transporte y traslado).

Por otro lado, al iniciar el proyecto se midió la diferencia que existía entre la cantidad de ventas y la cantidad de leads que se habían creado hasta el 08 de mayo de 2023, siendo 5795 ventas y 13793 leads. El día 04 de julio de 2023 se volvió a medir estos datos, obteniendo 6430 ventas y 14882 leads. Estos datos se traducen en la siguiente tabla:

⁶ Alto Valor Potencial.

⁷ Mediano Valor Potencial.

⁸ Bajo Valor Potencial.

Proyecto	Fecha	Ventas	Leads	Clientes/Leads
Sin proyecto	08/05/2023	5795	13793	42,01%
Con proyecto	04/07/2023	6430	14882	43,21%

Tabla 4: "Tabla comparativa situación con y sin proyecto del objetivo general"

Fuente tabla 4: "Creación propia a partir de base de datos Ventas_all y Leads"

En cuanto a los kpi de los objetivos específicos de la situación sin proyecto y la situación con proyecto, se obtuvieron los siguientes resultados:

Proyecto	Alcance campañas de correos (F2)	Interacción campañas de correos (F3)	Ventas Cross-Selling (F4)
Sin proyecto	38,80%	1,90%	0
Con proyecto	50,60%	3%	25

Tabla 5: "Tabla comparativa situación con y sin proyecto de objetivos específicos"

Fuente tabla 5: "Creación propia a partir de los resultados obtenidos"

En la Tabla 5 se consideran los datos de la campaña de Cross-Selling para AVP como parámetros de la situación con proyecto y una campaña de marketing general que se realizó a toda la base de leads en abril de 2023 para la situación sin proyecto.

b. Impacto.

Analizando el impacto de la situación con proyecto por objetivo general y objetivos específicos se observa que:

- Según el kpi del objetivo general (F1) que mide la proporción de ventas por leads, se puede decir que el impacto del proyecto fue positivo, ya que este kpi aumentó en 1.2 puntos porcentuales respecto a la situación sin proyecto. Si bien se generó un aumento significativo en este valor, no se alcanzó el valor de 1.5 puntos porcentuales definido en el objetivo general.
- El primer kpi de los objetivos específicos (F2) que mide el porcentaje de apertura de emails de la campaña de correos obtuvo un impacto positivo si se compara con la campaña de Cross-Selling para AVP que es la campaña principal de acuerdo al proyecto, logrando superar el objetivo de alcanzar un 50% de correos abiertos.
- Del mismo modo, el segundo kpi de los objetivos específicos (F3) que mide el porcentaje de interacción en los emails abiertos, se compara con la campaña de Cross-Selling para AVP, obteniendo un 3% de interacciones, valor superior al original de 1.9% pero no lo suficiente para cumplir con el objetivo de 4% establecido

anteriormente, por lo que se considera positivo respecto al valor original pero negativo respecto al objetivo.

- Finalmente, el tercer kpi de los objetivos específicos (F4) que mide la cantidad de ventas realizadas a partir de la campaña de Cross-Selling. Se considera que el impacto en este kpi es positivo, ya que se supera el objetivo establecido de 15 ventas realizadas a partir de esta campaña. Si bien estas 25 ventas todavía se encuentran en proceso, por factor de tiempo se decide considerar los clientes que están en conversación con los KAM como ventas.

Luego se mide el impacto que se genera en los ámbitos económico, social y medioambiental:

- Económico: si bien no es posible actualmente tener una evaluación económica precisa por no contar con los resultados finales, se hace una estimación contando como venta a la gente que se encuentra realizando una cotización y se cuenta como valor por venta 1.6983 UF, valor estimado por venta calculado anteriormente en los objetivos.

CAC	Costo humano mensual	Ventas Cross-Selling	Valor en UF	TIO	Inversión	Ventas Mayo	Ventas Junio	VAN	TIR
3	250,21	25	42,45	0,988	-4017,63	30607,63	31031,14	57084,75	752%

Tabla 6: "Tabla de evaluación económica"

Fuente tabla 6: "Creación propia a partir de los resultados obtenidos"

De esto, se puede decir que tanto el VAN como la TIR son positivas, por lo que se le recomienda a la empresa invertir en este proyecto.

- Social: Este proyecto no tiene mayor efecto en el aspecto social, sin embargo, si se llevan a cabo ciertas consideraciones con los clientes de alto valor potencial, se puede mejorar la imagen de la empresa al ofrecer un mejor servicio al cliente, más allá de esto, no se encuentra ningún impacto en lo social.
- Medioambiental: El impacto medioambiental que genera el proyecto no es muy grande pero al reducir la cantidad de leads que se generan en vano, se produce un menor impacto negativo por la contaminación de datos que se genera por el almacenamiento de grandes cantidades de información.

El impacto generado en los procesos y en los trabajadores es levemente positivo. Por el lado de los procesos, no se deben realizar acciones extra en ninguna de las áreas, por lo cual no afecta realmente a cómo se viene haciendo los procesos, pero se agrega una puntuación a las bases de datos que permiten identificar más fácilmente la prioridad de atención de los clientes. Mientras que por el lado de los trabajadores, sus funciones seguirán siendo las mismas pero perderán menos tiempo atendiendo clientes que no otorguen valor a la empresa.

Si bien se le dará prioridad a unos clientes antes que a otros, no se perjudica a ninguna persona u organización, por lo que no se incurre en dilemas éticos que puedan afectar a la empresa.

8. Conclusión:

En conclusión, es posible predecir si un cliente comprará un seguro o no hasta cierto grado si se consideran las variables indicadas, en este caso se utilizaron como variables el giro, compañía, comuna, producto, grupo de producto y el estado del cliente, aunque se estima que la predicción podría haber sido más precisa si se contara con la información del monto asegurado y una menor cantidad de errores en la base de datos. Además, con una buena predicción se pueden realizar distintas acciones de ventas y marketing que ayuden a realizar más ventas por un costo de adquisición de clientes menor. De estos datos se pueden obtener patrones y otros datos que ayuden en la generación de estrategias de ventas para futuros, actuales y antiguos clientes.

En cuanto al objetivo general, se concluye que si bien el objetivo establecido no se cumplió en el tiempo establecido, hay evidencia para creer que en cuanto se termine la campaña de Cross-Selling, este valor será alcanzado y posiblemente superado, actualmente el aumento es de 1.2 puntos porcentuales, 0.3 puntos porcentuales más bajo de lo estimado.

Por otro lado, en cuanto a los objetivos específicos, se obtuvo que un 50.6% de los mails fueron abiertos en las primeras 24 horas y un 3% de ellos obtuvieron una interacción con el link para agendar hora. Estos valores no son los resultados esperados según los kpi establecidos, pero significan una mejora para la empresa de todas maneras. En cuanto a la cantidad de ventas realizadas por la campaña de Cross-Selling, como se pospuso el lanzamiento de la campaña, no hubo tiempo suficiente para concretar ninguna de las ventas, por lo que momentáneamente se utilizará como kpi la cantidad de cotizaciones y contactos que se hicieron a partir de esta campaña, obteniendo 25 cotizaciones, número no menor que supera el objetivo que se estableció en los objetivos específicos.

Por lo que se puede concluir que si bien los objetivos general y específicos no alcanzaron las metas establecidas anteriormente del todo, se consiguieron avances que generaron ganancias a la empresa y a la larga promete entregar mejores resultados que si no se aplicara el proyecto.

9. Discusión:

Los mayores logros obtenidos en este proyecto son:

- Valor potencial: Sistema que permite agregar valor a los clientes de acuerdo a la probabilidad que representan de realizar una compra. Este sistema permite priorizar a los clientes que realmente se encuentran interesados de comprar un producto con aquellos que solo se encuentran probando la página o simplemente no están interesados en comprar pólizas de seguros, de esta manera se reduce considerablemente los gastos por adquisición de clientes y se concentran los esfuerzos de ventas en los verdaderos potenciales clientes.
- Cross-Selling: Campaña de Marketing que busca vender nuevos productos a clientes que cuenten con uno o más productos de Simplee basados en el rubro de la empresa (giro) y en el producto contratado anteriormente. Esta campaña permite sacar un mayor provecho de los clientes de Simplee, mejorar la fidelización de los clientes y conseguir nuevas ventas sin incurrir en nuevos gastos.

Para realizar este proyecto se debieron superar varias dificultades producidas por distintos factores:

- Por el lado técnico, hubo algunas dificultades en el tratamiento de las bases de datos y al modelar el árbol de decisiones debido a la capacidad del computador donde se realizó.
- Por el lado del factor humano, hubieron algunos retrasos en la planificación del proyecto debido a que el equipo del departamento de TI de la empresa se encontraba constantemente ocupado y se demoró en entregar las bases de datos necesarias para el proyecto, además, debido a pedidos del jefe se debió retrasar el envío de la campaña de Cross-Selling un mes, por lo que resultó imposible obtener mayores resultados.
- Por el lado del factor tiempo, los retrasos ocasionados por el factor humano no permitieron obtener los resultados completos del proyecto, ya que el flujo de la campaña de Cross-Selling requiere de 14 días entre correos.

El proyecto en sí quedó listo para implementarse y se puso a prueba la primera etapa de Cross-Selling. No obstante, para obtener resultados completos que evalúen correctamente el funcionamiento del proyecto se debe completar el flujo de esta campaña, lo que tomaría 28 días más al menos para terminar de enviar los correos y luego unos días para concretar las ventas, ya que primero se debe realizar la cotización de las pólizas solicitadas.

Se recomienda separar a los futuros clientes en distintos segmentos según su valor potencial para manejar los gastos en marketing de acuerdo al potencial de convertirse en una venta. Al mismo tiempo se deben realizar más testeos del funcionamiento de este y evaluar solicitar datos más relevantes para este estudio a los clientes para luego realizar una predicción de valor potencial más certera y manejar mejor los esfuerzos del equipo de ventas.

10. Referencias:

- [1] Diego Gonzalo Trujillo Vereau. (2021). *Calidad del servicio y fidelización de los clientes en la empresa Asesores y corredores de seguros Riley Eirl*, 2021. Obtenido de <https://repositorio.upn.edu.pe/bitstream/handle/11537/30307/Tesis.pdf?sequence=1&isAllowed=y>
- [2] Pía Cartes Henríquez. (2015). *Diseño de un modelo integral para la gestión de la retención de clientes de una corredora de seguros*. Obtenido de <https://repositorio.uchile.cl/bitstream/handle/2250/132071/Diseno-de-un-modelo-integral-par-a-la-gestion-de-la.pdf?sequence=4&isAllowed=y>
- [3] Paola V. Avellaneda C., Arturo Valencia S. (2015). *Propuesta de Mejora en el Proceso de Emisión de Pólizas y Servicio Posventa en una Empresa de Corredores de Seguros*. Obtenido de https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/620544/Avellaneda_cp.pdf?sequence=1
- [4] Luz de los Angeles M. Asencio D., Ricardo H. Chiang C., Fernanda L. Crisóstomo F., Gisela V. Hernández Q., Almendra S. Lajo A. (2021). *Técnicas de Machine Learning para la clasificación automática de clientes en una empresa de seguros*. Obtenido de https://repositorio.esan.edu.pe/bitstream/handle/20.500.12640/2933/2021_IIC_21-2_07_TC.pdf?sequence=1&isAllowed=y
- [5] Daniela González, Néstor Rincón. (2021). *Ejecutar un modelo de Machine Learning para identificar los clientes potenciales basados en un proceso probabilístico para la empresa Dell Technologies*. Obtenido de <https://repository.unimilitar.edu.co/bitstream/handle/10654/38562/GonzalezGordilloDanielaRinconManjaresNestor2021.pdf.pdf?sequence=1&isAllowed=y>
- [6] Álvaro A. Forero G. (2021). *Machine Learning para la segmentación y optimización de los costos de adquisición de clientes*. Obtenido de https://repository.libertadores.edu.co/bitstream/handle/11371/4971/%c3%81LVARO_ANTONIO_FORERO_GONZ%c3%81LEZ_2021.pdf?sequence=1&isAllowed=y

11. Anexos:

Anexo A.

Dato extra para el contexto de la empresa que ayuda a entender mejor cómo funciona la empresa y los productos que ofrece. Los tipos de seguros que Simplee ofrece son los siguientes:

- **Responsabilidad Civil:** El Seguro de Responsabilidad Civil es una póliza que cubre a la pyme frente a eventuales demandas o reclamos formales que efectúen terceros como consecuencia de algún daño físico o material que hayan sufrido producto a un evento causado por tu empresa.
- **Accidentes Personales:** El Seguro de Accidentes Personales es una póliza que protege a los trabajadores asegurados de una empresa y entrega un respaldo económico a su familia en caso de muerte accidental.
- **Incendios, Sismo y Robo:** En caso de poseer o arrendar una propiedad (oficina, tienda, local comercial, almacén, bodega, fabrica, etc.), el seguro de Incendios, Sismos y Robo asegura hasta 100% de la infraestructura y el contenido del activo comercial del solicitante de la póliza de seguro.
- **Colectivo de Salud:** Los seguros complementarios de salud cubren todas las prestaciones médicas después de Isapre y Fonasa. Asegura a partir de 5 trabajadores con estas pólizas especialmente diseñadas para pymes.
- **Garantía:** Es solicitado como requisito en licitaciones. En caso de que no se cumpla el contrato por diferentes motivos, el cliente cobrará directamente el monto asegurado a la compañía de seguros.
- **Transporte y Carga:** Asegura la carga que lleva en el medio transportador. Existen 2 modalidades: póliza flotante o flat (solo carga, no importan los datos del medio transportador) y responsabilidad máxima (solo carga pero con datos del medio transportador).

Anexo B.

Para entender mejor cómo es el proceso y las posibles razones de la pérdida de clientes, se hizo un diagrama que muestra cómo es la interacción del cliente con la empresa:

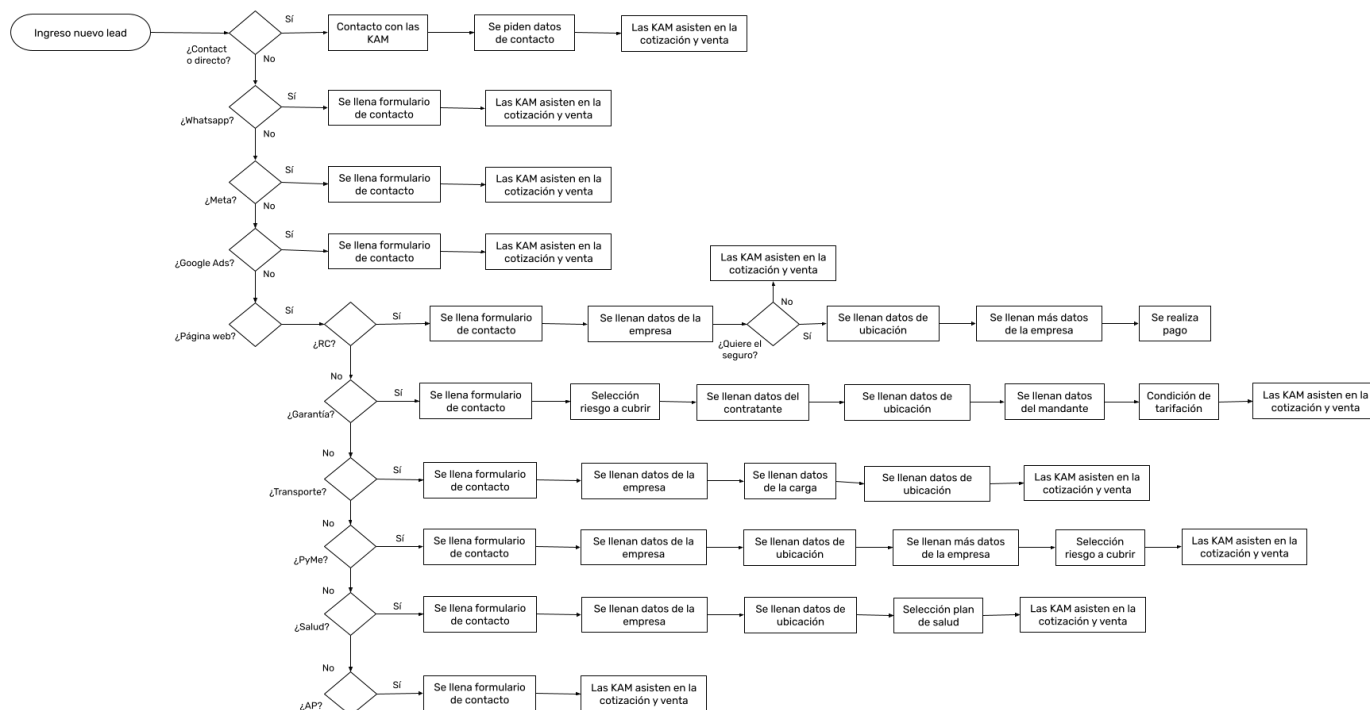


Figura 1: “Diagrama de flujo del proceso que viven los clientes al cotizar seguros con Simplee”

Fuente figura 1: “Creación propia a partir de flujo de experiencia del cliente entregado por Simplee”

Este diagrama se traduce como: Primero un posible lead escucha/lee de Simplee a través de anuncios de Google Ads, una de las fuentes de Meta, Whatsapp, la página web de Simplee o por contacto directo con otra persona. La persona ingresa al formulario de contacto de alguna de estas fuentes, lo rellena, se crea el Lead en el sistema para que finalmente alguno de los KAM⁹ se contacte con este Lead y pueda asistirlo con cualquier información que requiera sobre el o los productos solicitados. El Lead queda asignado con un KAM en específico que lo ayudará con cualquier información que requiera, realizará la cotización para encontrar el seguro que más le convenga y si finalmente este desea hacer la compra, se realiza el cobro debido en la cantidad de cuotas que prefiera y con el medio de pago que le resulte conveniente, ya sea con débito o crédito.

⁹ KAM: Key Account Manager, gestor de cuentas de la empresa, trabaja en el departamento de ventas.

Por otro lado, si el Lead ingresa por la página web, el proceso será un poco diferente del resto. En primer lugar, si el cliente desea comprar un seguro de Responsabilidad Civil (RC), este se puede realizar 100% por la página web y en el instante. Primero se llena el formulario de contacto al igual que en todos los casos, luego se llenan los datos de la empresa, luego dependiendo de si desea continuar con la compra del seguro a través de la página o si requiere más información o simplemente prefiere ser contactado, el Lead es contactado por un KAM que se le asignará o podrá finalizar el proceso por la página. Si decide terminar el proceso online, se le pide llenar los datos de ubicación de la empresa, luego se pide llenar unos datos más de la empresa y finalmente se procede a realizar el pago para finalizar la venta.

Para los otros tipos de productos, el proceso es bastante similar entre sí, variando en los datos que se piden, pero con la diferencia de que estos deben de finalizar la compra conversando con el KAM asignado y no a través de la página web. En general, se piden los datos de contacto, datos de ubicación y algunos datos de la empresa que ayuden a la cotización para finalmente ser contactado por uno de los KAM y finalizar la venta.

Este proceso de recolección de datos se debe realizar cada vez que un Lead desee cotizar un nuevo producto y requiere mucho tiempo realizar todo este proceso, por lo que puede llegar a resultar bastante tedioso tanto para los clientes como para los KAM que realizan las ventas.

Anexo C.

En cuanto a las herramientas y/o métodos existentes en la industria o literatura que sirven para solucionar problemas de fidelización de clientes e incremento de ventas, se realizó un estudio en profundidad de diferentes alternativas.

En primer lugar se realiza una investigación de la influencia que tiene la calidad del servicio al cliente en la fidelización de los clientes, evaluando distintos aspectos que forman una buena calidad de servicio.

Se concluye que sí existe una relación positiva entre la dimensión fiabilidad y la variable fidelización de los clientes, aunque en un nivel bajo; ello quiere decir que la empresa debe transmitir más confianza a sus clientes y brindar los servicios con precisión para que el cliente se sienta confiable con el servicio y sienta compromiso con la empresa (Trujillo Vereau, 2021).

Se concluye que existe relación positiva entre la capacidad de respuesta y la fidelización de los clientes, pero en un nivel bajo; dando a entender que los factores como la rapidez del servicio, solución inmediata y atención a reclamos no están incidiendo considerablemente en la lealtad de las personas por la empresa (Trujillo Vereau, 2021).

Se concluye que sí existe relación positiva moderada entre la dimensión seguridad y la variable fidelización de los clientes; de manera que los clientes perciben como agradable la atención de los colaboradores y la credibilidad del servicio que se ofrece; por ello se sienten satisfechos y presentan cierto grado de fidelidad hacia la empresa (Trujillo Vereau, 2021).

Se concluye confirmándose que existe relación positiva entre la tangibilidad y la fidelización de los clientes en un nivel bajo; dando a entender los aspectos de las instalaciones, infraestructura, equipos y apariencia del personal no repercuten significativamente en la fidelidad de las personas por la empresa (Trujillo Vereau, 2021).

Se concluye confirmándose que existe relación positiva entre la empatía y la fidelización de los clientes en un grado débil, demostrando que los factores como la atención personalizada, la comprensión de las necesidades de los clientes y comunicación con ellos no están repercutiendo como se espera en la fidelidad de los clientes (Trujillo Vereau, 2021).

Se concluye que sí existe relación positiva y considerable entre la calidad del servicio y la fidelización del cliente; de manera que brindar un servicio confiable, con seguridad, teniendo interés y preocupación por el cliente influye significativamente en su fidelidad hacia la empresa (Trujillo Vereau, 2021).

Por otro lado, se encontró una propuesta de modelo integral para la retención de clientes que demuestra que por muchos factores conviene más retener a un cliente que realizar una nueva venta, por lo que es importante poner énfasis en realizar buenos planes de retención.

Este artículo indica que más allá de hacer una simple segmentación, se debe hacer foco en distintos factores e indica que los primeros 5 meses de vigencia son cruciales para la retención de clientes.

La estrategia planteada en este artículo busca llenar un vacío importante de servicio que es la comunicación con el cliente, mejorando la experiencia del cliente. Se recomienda hacer un ciclo de información para los segmentos de clientes más masivos para que los clientes conozcan con claridad el seguro que están pagando y los beneficios que este tiene tanto por la cobertura que le proporciona como por los beneficios adicionales que la mayoría de los productos poseen, de modo que los clientes valoren lo que tienen y se quieran quedar en la empresa.

Esta idea se basa en la información recopilada para entregar un mejor servicio e información a los clientes por segmento, de esta manera mejora la experiencia del cliente y la retención de estos (Cartes Henríquez, 2015).

Además se estudió el posible impacto que podría tener la aplicación de modelos de predicción de Machine Learning para identificar, segmentar y sacar provecho en el rubro de las corredoras de seguros, concluyendo que las herramientas de Machine Learning tienen el potencial de generar eficiencias dentro de la empresa y fomentar el desarrollo de ventajas competitivas que impacten positivamente en el desempeño general de la compañía. La empresa puede desarrollar una estrategia de marketing personalizado utilizando los resultados del modelo propuesto, logrando un mayor nivel de ventas y fidelización por parte de los clientes (Asencio D. et al., 2021).

Otro uso de Machine Learning puede ser para predecir el valor potencial que tienen los clientes, es decir, la probabilidad de que una persona realice una compra de uno de los productos ofrecidos por la empresa. Se cree que es importante fomentar la utilización de herramientas que permitan acercarse a las cuentas prospecto, esto permitirá forjar relaciones entre vendedores y empresas y con ello la identificación de oportunidades que con el manejo correcto pueden convertirse en ventas. Utilizar listas específicas de clientes a atacar permite un trabajo asertivo y conversaciones que generen valor (Gonzalez & Rincón, 2021).

Como última función de Machine Learning se estudió el efecto que esto puede tener para reducir los costos de adquisición de clientes. De esto se obtiene que realizar una clusterización previo a una regresión logística entrega un mejor resultado, además que al identificar los posibles clientes interesados (valor potencial), se reduce el esfuerzo y gasto realizado en contactar y atraer leads sin considerar si es un potencial cliente o no (Forero, 2021).

Otra propuesta para mejorar la experiencia de los clientes se recomienda implementar una plataforma web para que los clientes puedan consultar información relacionada a sus

pólizas. Con ello, se favorecerá la autogestión, reduciendo el trabajo del equipo, habilitando a ocupar el tiempo obtenido en otras áreas del negocio y brindando información al instante a los consumidores (Avellaneda C. & Valencia S., 2015).

Anexo D.

Definición de las alternativas de solución analizadas en la tabla de adecuación y utilidad en el marco teórico.

- **Agregación de valor al producto ofrecido:** Esta alternativa de solución busca fortalecer la fidelización y hacer los productos ofrecidos más atractivos para los leads mejorando la experiencia de producto, lo que se vería reflejado como un aumento de las ventas tanto para antiguos clientes (Upselling) como para nuevos clientes. Para esto, se debe mantener un conocimiento actualizado de los productos que se ofrecen en el mercado, que los socios de la empresa ofrecen y los productos que los posibles clientes requieren (los seguros que las PYMEs, grandes empresas y personas naturales requieren y utilizan actualmente). De esta manera se podrá ofrecer opciones de seguros más adecuadas y competitivas a los clientes, además de ayudar a identificar oportunidades para upselling o cross-selling, lo que podría aumentar los ingresos y comisiones de la empresa. Para esto, se debe hacer un análisis tanto de los datos de Simplee como de otras empresas que compitan en el mercado de las corredoras de seguros en búsqueda de patrones en cuanto a los productos más vendidos y toda la información relevante para la cotización y venta de una póliza (datos como la ubicación, tamaño de las empresas, rubro de la empresa, entre otros). A partir de esta información se pueden desarrollar distintos planes de venta de los productos investigados, realizar acciones de Upselling y Cross-Selling, negociar nuevas ofertas o servicios con las empresas asociadas de seguros y buscar una manera de hacer más atractivo el producto.
- **Identificación de leads por valor potencial:** Esta alternativa busca crear un sistema de segmentación que permita distribuir de mejor manera los esfuerzos y recursos de la empresa y de esta manera reducir los costos de adquisición de clientes y crear planes de acción de ventas/marketing más específicos, buscando atraer a aquellos leads que por X razón no resultó en una venta y a aquellos que ya cuentan o contaron con uno o más productos de Simplee. Para esto se debe realizar un modelo de machine learning que aprenda de la data recolectada por la empresa, esto permite hacer un sistema que permita predecir qué tan probable es que cada lead/cliente realice una compra de uno o varios productos de la empresa. Luego, se debe buscar patrones sobre los productos y giro de las empresas que tienen una mayor probabilidad de convertirse en clientes. A partir de esto se realizan flujos que ayuden a ofrecer los productos que, según la información obtenida, será lo que posiblemente el cliente requiera o desee comprar de acuerdo a cada caso en específico.

- **Simplificación de flujo de interacciones:** Esta alternativa busca mejorar la experiencia de compra y servicio de los clientes/leads, con tal de mejorar la fidelidad de los clientes y reducir la brecha de usuarios que inician el proceso de compra pero finalmente no realizan la compra. Para esto se deben hacer cambios en la página, empezando por agregar una opción de usuario para almacenar toda la información de los leads para que este no deba ingresar todos sus datos cada vez que quiera cotizar un producto. Además, esto permitirá centralizar toda su información para que este pueda acceder a su información y a su cartera de productos cada vez que quiera. Esto hace que el proceso sea menos tedioso para el cliente, le da acceso a la información que requiera y reduce el trabajo que debe realizar el equipo de Simplee por cada cliente. Por otro lado, esto requiere que se cree una plantilla como diseño de la nueva página web, definir cada función e interacción en la página, trabajar junto con TI en el Backend y Frontend de esta, adaptar los otros procesos al nuevo formato de la página, crear un nuevo flujo de la experiencia del cliente y capacitar al equipo completo al uso de la nueva página web.

Anexo E.

En el Anexo E se entra en mayor detalle de los procesos realizados en el desarrollo de la solución para el problema planteado. Se explica paso a paso el manejo de la base de datos, las decisiones tomadas, los análisis que se realizaron y la implementación de la solución.

En primer lugar se examinan las distintas bases de datos con las que cuentan en Simplee en busca de la base de datos que entregue información más relevante para realizar una predicción de valor potencial de clientes. A partir de lo investigado sobre el mercado de corredoras de seguros a través de distintos sitios web, consultas en Chat GPT y recomendaciones de parte de distintos trabajadores de la empresa, se llegó a la conclusión de que las variables que tienen un mayor efecto en la venta de una póliza de seguros son: Comuna, Giro, Compañía, Producto, Grupo de producto, Estado del cliente y Monto asegurado. Considerando esto y que es la base de datos con menor cantidad de errores o datos nulos, se decide proseguir con la base de datos de Ventas_all.

Antes de comenzar a trabajar en Python de Jupyter Lab, se pasó la base de datos a Power BI donde se procedió a arreglar los tipos de variables de las variables “Fecha de contacto”, “Fecha de emisión”, “Fecha de inicio de vigencia” y “Fecha de Terminación de vigencia” a “datetime64[ns]” y las variables “RUT” y “Telefono” a “object”. Además se procedió a eliminar otras variables que no entregan ningún valor a la predicción ni a la empresa, reduciendo la base de datos a 27 columnas y 5700 filas.

Al comenzar a trabajar con la base de datos se encontraron algunos problemas que podrían complicar la calidad de la predicción. En primer lugar, la base de datos tiene muchos datos vacíos y errores en como están escritas los valores de cada variable, por ejemplo, en la variable comuna existe el valor “Las Condes”, el valor “LAS CONDES”, el valor “Las Condes, Santiago” y más como estos que le quitan valor a esta al segmentar el valor “Las Condes” en cada variación de cómo se escribió esta comuna. Por otro lado, otra complicación es que de las 27 variables existentes, solo 3 de ellos son valores numéricos y las demás son valores categóricos, lo cual condiciona a utilizar como variable independiente a una de estas 3 variables y a que no se pueda utilizar cualquier modelo de Machine Learning para esto. Además, al ser en su mayoría valores categóricos, se deben realizar otras acciones extra antes de poder realizar un cluster para esta data.

Se comienza a trabajar en Python, mostrando la información de la tabla (columnas, valores no nulos y tipo de dato), luego se estandarizan las variables numéricas que se tienen para graficar en histogramas y revisar si estas presenta una forma de campana de Gauss y definir si presenta un comportamiento normal. Para proseguir, se decide definir a la variable “Comisión Final” como la variable independiente.

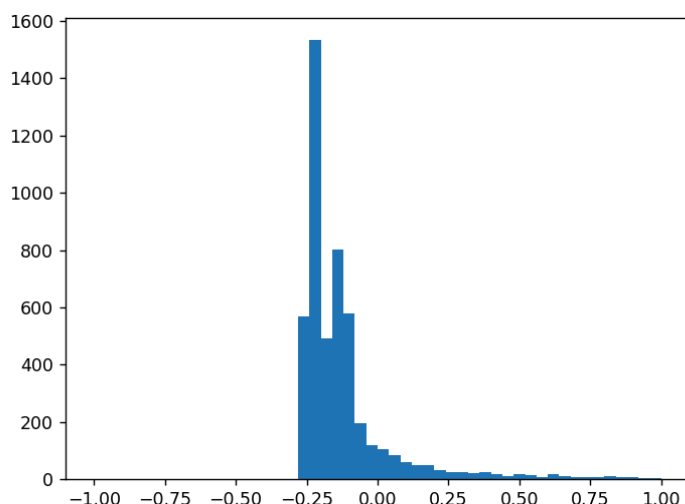


Figura 6: “Histograma de la variable Comisión Final estandarizada que demuestra un comportamiento normal”

Fuente figura 6: “Creación propia a partir de la base de datos de ventas de Simplee con Python de Jupyter Lab”

Luego de confirmar que la variable independiente presenta una distribución normal, se puede proseguir con el modelo de machine learning, ya que si una variable no sigue una distribución normal, los resultados y conclusiones pueden ser incorrectos o sesgados.

A continuación se realizan los análisis estadísticos pertinentes para evaluar a la variable independiente.

Promedio	3.452710705886525
Moda	1.6983
Mediana	1.37
Cuartil 1	0.59
Cuartil 2	1.37
Cuartil 3	2.2425
Cuartil 99%	46.734799999999978
Varianza	158.3090911300331

SD	12.58209406776285
Máximo	432.0
Mínimo	0.0
Rango	432.0
IQR	1.6525000000000003
Asimetría	15.408482346282314
Kurtosis	366.83032895626775

Tabla 7: "Tabla de análisis estadísticos"

Fuente tabla 7: "Creación propia a partir de datos de la base de datos Ventas_all analizados en Python"

- Promedio: El promedio de la variable "Comision Final" es aproximadamente 3.45.
- Moda: La moda de la variable es 1.6983, lo que indica que ese valor es el más frecuente en los datos.
- Mediana: La mediana de la variable es 1.37. Esto significa que el 50% de los valores se encuentran por encima de este valor y el otro 50% se encuentra por debajo.
- Cuartiles: Los cuartiles dividen los datos en cuatro partes iguales. El primer cuartil (Q1) es 0.59, el segundo cuartil (Q2) es igual a la mediana (1.37) y el tercer cuartil (Q3) es 2.2425.
- Cuartil 99%: El cuartil 99% es 46.7348, lo que indica que el 99% de los valores de "Comision Final" se encuentran por debajo de este valor.
- Varianza: La varianza de la variable es aproximadamente 158.31.
- Desviación estándar: La desviación estándar de la variable es aproximadamente 12.58. Indica la dispersión de los datos con respecto a la media.
- Máximo: El valor máximo de la variable es 432.0.
- Mínimo: El valor mínimo de la variable es 0.0.
- Rango: El rango es la diferencia entre el valor máximo y el valor mínimo. En este caso, el rango es 432.0.
- IQR: El rango intercuartílico (IQR) es la diferencia entre el tercer cuartil (Q3) y el primer cuartil (Q1). En este caso, el IQR es 1.6525.
- Asimetría: La asimetría de la variable es 15.41. Indica la falta de simetría en la distribución de los datos.
- Kurtosis: La kurtosis de la variable es 366.83. Indica el grado de apuntamiento o

aplastamiento de la distribución en comparación con una distribución normal.

Luego se procede a hacer un gráfico de distribución de Kernel como otro método para revisar la distribución de la variable “Comision Final”.

Gráfico de Densidad de Kernel para la variable Comision Final

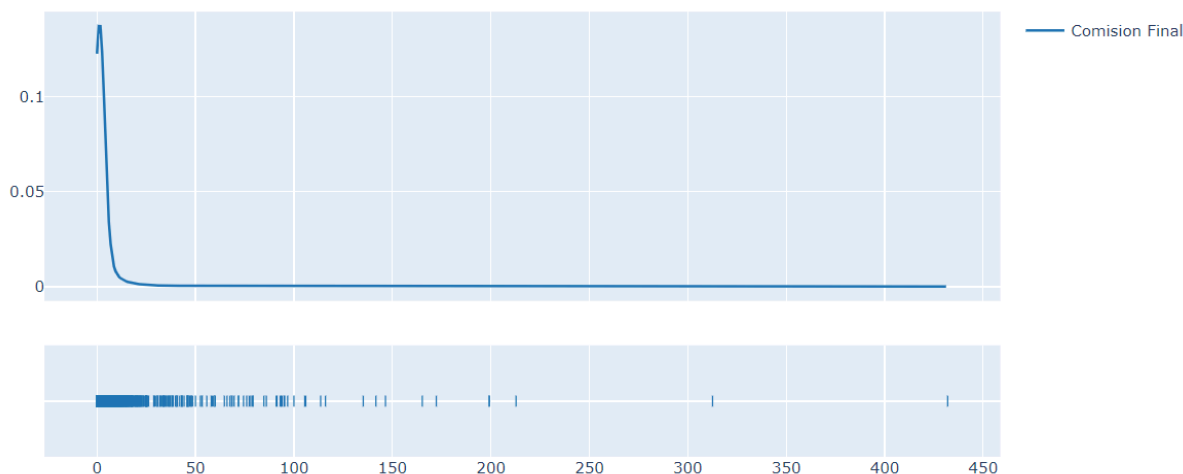


Figura 7: “Gráfico de Densidad de Kernel para la variable Comision Final”

Fuente figura 7: “Creación propia a partir de datos de la base de datos Ventas_all analizados en Python”

Continuando con la limpieza de la base de datos, se procede a identificar y eliminar los datos outliers de la base que posteriormente servirá para predecir el valor potencial de los clientes. Para esto se utiliza el Local Outlier Factor (LOF) en las variables numéricas “Prima” y “Comision Final” luego de haber eliminado los datos nulos para que LOF no tenga errores. Luego de varias iteraciones se encontró que el número indicado de vecinos que se deben considerar para esto, tiene que ser de 5 vecinos, ya que es donde la curva se suaviza y la diferencia de vecinos deja de cambiar mucho la cantidad de outliers encontrados.

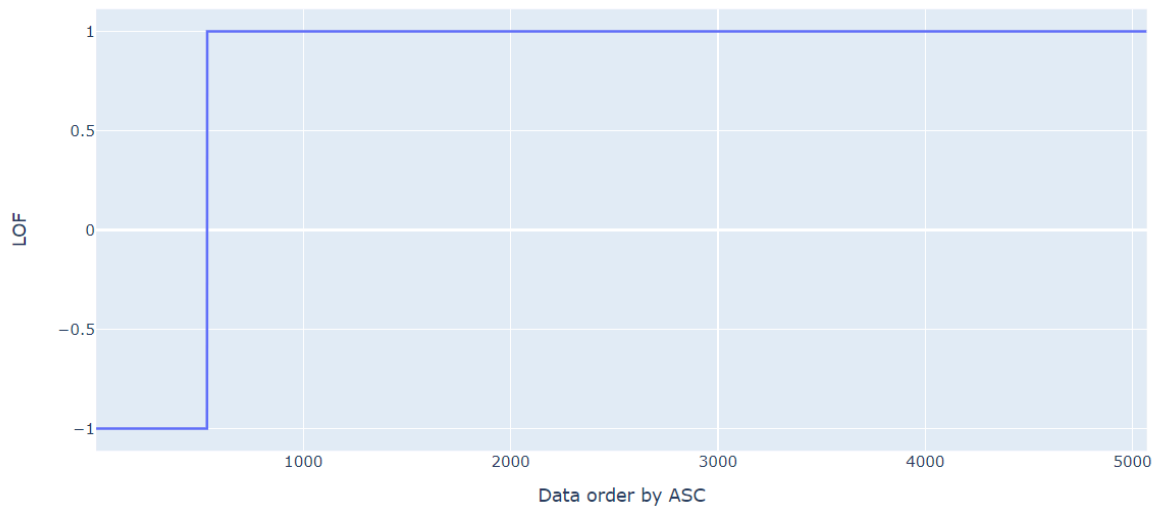


Figura 8: “Gráfico de LOF que muestra la cantidad de datos outliers que se encuentran en la base de datos, siendo los datos con LOF de -1 considerados datos anómalos (outliers)”

Fuente figura 8: “Creación propia a partir de datos de la base de datos Ventas_all analizados en Python”

De esto se obtiene que 536 datos son considerados outliers con un LOF igual a -1, por lo que se procede a eliminar estos datos de la base de datos, quedando con una base de datos de 4533 datos.

Luego, se procede a utilizar el código de LabelEncoder para darle un valor numérico a las variables categóricas y poder comenzar a crear los clusters. Para esto se realiza un gráfico para verificar cuántos clusters debemos crear.

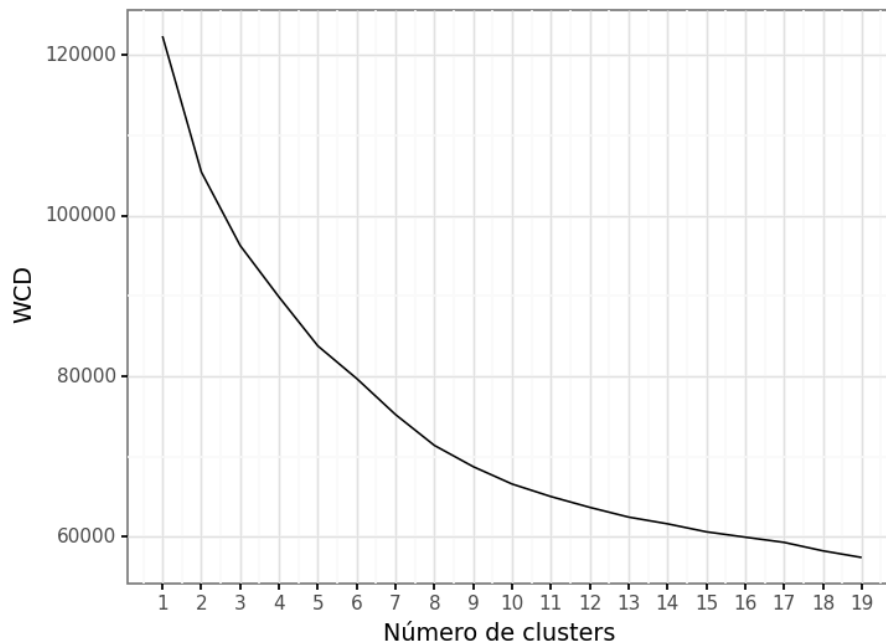


Figura 9: “Gráfico que muestra la cantidad de clusters que se deben crear para obtener una clusterización exitosa”

Fuente figura 9: “Creación propia a partir de datos de la base de datos Ventas_all analizados en Python”

De este gráfico se obtiene que el número total de clusters debe de ser de 2, siendo donde la curva comienza a estabilizarse. En el cluster 0 quedan 2852 datos y en el cluster 1 quedan 1681 datos.

Una vez creados los clusters, se realiza un análisis de datos sesgados, ya que antes de generar el modelo es importante entender todas las variables. Se observa la existencia de diferentes proporciones en algunas variables, si esto sucede, se podría tener un problema GIGO (Garbage In, Garbage Out). Esto es muy importante porque los modelos de Machine Learning ayudan a las personas a tomar decisiones importantes y si el modelo se encuentra sesgado, esto hará que se propague en diferentes áreas de la sociedad. Por suerte, no se encontraron valores significativos en ninguna de las variables que indiquen un problema de GIGO.

Finalmente, con todos los análisis ya realizados, se puede proceder a escoger el modelo de Machine Learning que se utilizará como solución a este problema. Como se trata de una base de datos de valores categóricos, el modelo de Machine Learning que mejor resuelve el problema es Árboles de decisiones, ya que este modelo puede manejar de mejor manera las variables categóricas, permite una interpretación y comprensión mayor al ser representado visualmente, permite identificar las características más relevantes para la toma de decisiones, permite captar patrones complejos en los datos, muestra una mejor

resistencia al ruido y valores atípicos (especialmente importante para esta base de datos) y tiene una excelente eficiencia en tiempo de entrenamiento y predicción.

Se seleccionan las variables dependientes más importantes que se utilizarán para realizar este modelo ("Compañía", "Giro", "Comuna", "Estado", "Grupo de Producto", "Producto", "Comisión Final", "Prima") y la variable que se utilizará como variable independiente ("etiquetas") que corresponde a los cluster definidos anteriormente. Luego se crea un grupo de testeo y otro de entrenamiento, que corresponden a un 30% y 70% de la muestra respectivamente.

Para llevar a cabo este modelo se utiliza el código "DecisionTreeClassifier" y se verifica los puntajes del grupo de testeo y de entrenamiento, obteniendo los siguientes valores:

- score en data de entrenamiento: 0.998
- score en data de validación: 0.846

Esto indica que se tiene overfitting. El modelo tiene una performance excelente en el dataset de entrenamiento, sin embargo, no generaliza bien para los datos de testing. Por esto, se realizan los ajustes adecuados de "Max depth", obteniendo los siguientes valores:

- train score: 0.8787
- test score: 0.8551

Ahora el problema de overfitting ya no afecta el modelo, por lo que se procede a graficar el árbol de decisiones y descargar la base de datos resultante de este proceso, donde se agrega una columna que indica a qué cluster pertenece cada dato e indica su valor al codificarse para poder interpretar el árbol de decisiones de manera más clara.

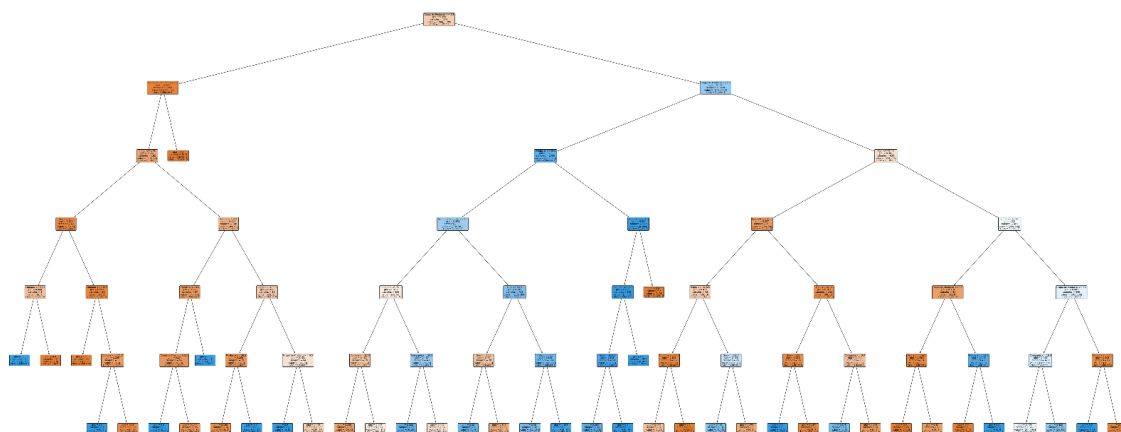


Figura 10: "Árbol de decisiones, bloques azules son del cluster 1 (potencial cliente) y bloques naranja son del cluster 0 (no potencial cliente)"

Fuente figura 10: "Creación propia a partir de la base de datos de ventas de Simplee con Python de Jupyter Lab"

Una vez obtenido el árbol de decisiones, se pasan las nuevas bases de datos mencionadas anteriormente a Power BI, se analiza la base de datos y se crean distintos filtros de acuerdo a los resultados que entrega el árbol de decisiones mostrado anteriormente. Gracias a estos filtros se pueden identificar los valores que debe de tener un cliente para ser considerado un cliente potencial.

Compañía

Compañía

assist_card

avla

bci

bice

bupa

cesce

chilena

chubb

consorcio

contempora

continental

hdi

inge

maxxa

orion

orsan

southbridge

suaval

sura

0

1

2

3

4

5

6

7

8

9

10

12

13

16

17

18

21

22

23

Comuna

Comuna

004SIQU 00

004SPEN 00

34

ANTOFAGASTA

Buín

CONCHALI

Cabrero

Calama

Calera de Tango

Cartagena

Castro

Cerrillos

Colchane

Colina

Concepción

Copiapó

Coquimbo

Curico

El Bosque

El Quisco

Estación Central

Huechuraba

1

3

9

11

20

21

23

24

26

28

30

33

38

39

41

45

46

50

53

56

57

61

Producto

Producto

ap_colectivo

ap_digital

ap_individual

ap_mas_vida_guardia

asiento_pasajero

asistencia_viaje

colectivo_salud

colectivo_vida

garantia

garantia_corfo

garantia_privada

garantia_publica

garantia_serotec

0

1

2

3

4

5

9

10

13

14

15

16

17

Giro

Giro

76.859.760-K

ABOGADOS

ACONDICIONAMIENTO DE EDIFICIOS. AIRE ACONDICIONADO.

ACTIVIDADES COMBINADAS DE SERVICIOS ADMINISTRATIVOS DE OFICINA

ACTIVIDADES DE AGENCIAS DE EMPLEO

ACTIVIDADES DE AGENTES Y CORREDORES DE SEGUROS

ACTIVIDADES DE CLINICAS VETERINARIAS

ACTIVIDADES DE CONSULTORIA DE GESTION

ACTIVIDADES DE CONSULTORIA DE INFORMATICA

ACTIVIDADES DE CONSULTORIA DE INFORMATICA Y DE GESTION DE INSTALACIONE

ACTIVIDADES DE CONSULTORIA DE INFORMATICA Y DE GESTION DE INSTALACIONES

ACTIVIDADES DE CONSULTORIA DE INFORMATICA Y DE GESTION DE

1

1

2

3

4

5

9

10

13

14

15

16

17

Estado

Estado

cliente_cerrado

cliente_no_aplica

derivado_a_kam

en_espera_de_confirmacion

en_espera_de_precio

no_contactar

8

0

1

2

3

5

6

Filtros

Filtros

Buscar

Buscar

Filtros de esta página

Filtros de esta página

Grupo de Produ...

Grupo de Produ...

es menor que 4

es menor que 4

Producto

Producto

es menor que 27

es menor que 27

Giro

Giro

es menor que 418

es menor que 418

Agregar campos de datos ...

Agregar campos de datos ...

Filtros de todas las páginas

Filtros de todas las páginas

Agregar campos de datos ...

Agregar campos de datos ...

Nombre	Apellido	Fecha de contacto	Correo	Campaña	Canal	Comuna	Telef®
Ezequiel	Vidal	28-04-2023 0:00:00	ezequiel@agenciafauna.com	Desde Clientes	Directo	Santiago	+569
Raúl	Córdova	28-04-2023 0:00:00	rcordovasilva@gmail.com	Desde Clientes	Directo	Copiapó	+569
David	Pérez	27-04-2023 0:00:00	pyppreparacionymantenencion@gmail.com	Conocido	Directo	La Florida	+569
Arturo	Acevedo	12-04-2023 0:00:00	arturo@arpiingenieria.cl	Referido del Cliente	Directo	Estación Central	+569
Natalia	Abuadba	31-03-2023 0:00:00	natalia.conservitda@gmail.com	Desde Clientes	Directo	San Bernardo	+569
Cristian	Zurita	30-03-2023 0:00:00	construccionesczs@gmail.com	Desde Clientes	Directo	Providencia	+569
Jordan	Aguilera	28-03-2023 0:00:00	j19.making@gmail.com	Seguro de Responsabilidad Civil Digital	Google Ads	La Serena	+569
Total							

Tabla 8: “Tablas que contienen los valores que un potencial cliente debe tener junto a los filtros creados a partir del árbol de decisiones”

Fuente tabla 8: “Creación propia a partir de los resultados obtenidos del trabajo en Python y diseñado en Power BI”

A partir de esto se generó una plantilla que entrega una puntuación de acuerdo a cuántas de sus variables cuentan con un valor que aparezca en la tabla mostrada anteriormente, siendo 6 el valor máximo (6 variables seleccionadas para la predicción) y 0 el valor mínimo. Para esto, a modo de prueba para ver la efectividad y para encontrar distintos patrones en los datos, se decidió dividir el valor potencial en 3 segmentos:

- Entre 5 y 6 puntos se considera Alto Valor Potencial (AVP).
- Entre 3 y 4 puntos se considera Mediano Valor Potencial (MVP).
- Entre 0 y 2 puntos se considera Bajo Valor Potencial (BVP).

	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK
1	Comision	Comision Fi	N Poliza	Fecha de en	Fecha de ini	Fecha de Te	Tipo de pag	Puntaje com	Puntaje G pi	Puntaje esta	Puntaje giro	Puntaje proxi	Puntaje corr	Total
2	1,6	1,6	560 12443074 0	2023-04-13	2023-04-13	2023-04-19	contado	1	1	1	1	1	1	6
3	68,62	68,62	W5340-007496	2023-04-26	2023-04-26	2024-01-30	contado	1	1	1	1	1	1	6
4	0,63	0,63	W1856-007497	2023-04-26	2023-04-26	2024-01-30	contado	1	1	1	1	1	1	6
5	0,63	0,63	W5030-007479	2023-04-26	2023-04-26	2024-01-30	contado	1	1	1	1	1	1	6
6	0,63	0,63	W2694-007478	2023-04-26	2023-04-26	2024-01-30	contado	1	1	1	1	1	1	6
7	0,63	0,63	W7042-007485	2023-04-26	2023-04-26	2024-01-30	contado	1	1	1	1	1	1	6
8	0,63	0,63	W7240-007503	2023-04-26	2023-04-26	2024-01-30	contado	1	1	1	1	1	1	6
9	0,63	0,63	W9645-007487	2023-04-26	2023-04-26	2024-01-30	contado	1	1	1	1	1	1	6
10	0,13	0,13	03-24-006537	2023-04-24	2023-04-30	2023-06-30	contado	1	1	1	1	1	1	6
11	0,13	0,13	03-24-004729	2023-04-24	2023-04-30	2023-06-30	contado	1	1	1	1	1	1	6
12	0,13	0,13	03-24-004941	2023-04-24	2023-04-30	2023-06-30	contado	1	1	1	1	1	1	6
13	0,3	0,3	222104955	2023-04-24	2023-04-30	2023-06-30	contado	1	1	1	1	1	1	6
14	0,3	0,3	222104961	2023-04-24	2023-04-30	2023-06-30	contado	1	1	1	1	1	1	6
15	0,3	0,3	222104975	2023-04-24	2023-04-30	2023-06-30	contado	1	1	1	1	1	1	6
16	0,3	0,3	221117210	2023-04-24	2023-04-30	2023-08-31	contado	1	1	1	1	1	1	6
17	1,74	1,74	223103792	2023-04-21	2023-03-28	2024-05-30	contado	1	1	1	1	1	1	6
18	0,92	0,92	31256	2023-04-28	2023-09-07	2024-01-31	contado	1	1	1	1	1	1	6
19	0,45	0,45	PENDIENTE	2023-04-28	2023-05-10	2024-04-26	contado	1	1	1	1	1	1	6
20	1,7	1,7	PENDIENTE	2023-04-28	2023-05-10	2024-03-18	contado	1	1	1	1	1	1	6
21	0,45	0,45	PENDIENTE	2023-04-28	2023-05-10	2024-04-26	contado	1	1	1	1	1	1	6
22	1,7	1,7	PENDIENTE	2023-04-28	2023-05-10	2024-03-18	contado	1	1	1	1	1	1	6
23	0,45	0,45	PENDIENTE	2023-04-28	2023-05-10	2024-04-26	contado	1	1	1	1	1	1	6
24	1,53	1,53	PENDIENTE	2023-04-28	2023-05-10	2024-03-18	contado	1	1	1	1	1	1	6
25	0,45	0,45	PENDIENTE	2023-04-28	2023-05-10	2024-04-26	contado	1	1	1	1	1	1	6
26	1,53	1,53	PENDIENTE	2023-04-28	2023-05-10	2024-03-18	contado	1	1	1	1	1	1	6

Tabla 9: “Plantilla de valor potencial que indica cuál es el valor potencial del cliente”

Fuente Tabla 9: “Creación propia a partir de la base de datos Ventas_all y los resultados obtenidos anteriormente en Power BI”

Finalmente, para implementar los resultados obtenidos, se crea un diagrama de flujo para cada segmento de valor potencial definido anteriormente basándose en los datos y patrones obtenidos al analizar los datos por segmento en Power BI. Estos flujos representan las acciones que se deben hacer en cada caso para realizar una campaña de Cross-Selling donde se saque provecho de los análisis realizados anteriormente y del valor potencial de los clientes.

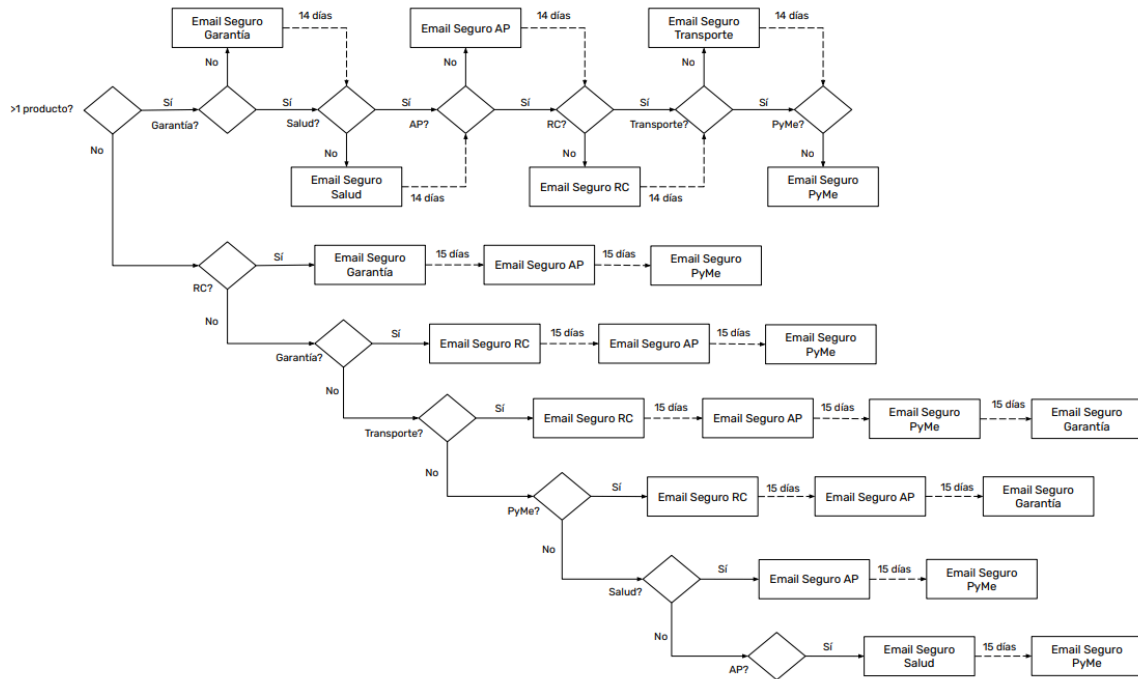


Figura 11: “Diagrama de flujo de Cross-Selling para segmento de alto valor potencial”

Fuente figura 11: “Creación propia a partir de patrones, datos obtenidos del árbol de decisiones elaborado anteriormente y ayuda del equipo de Marketing”

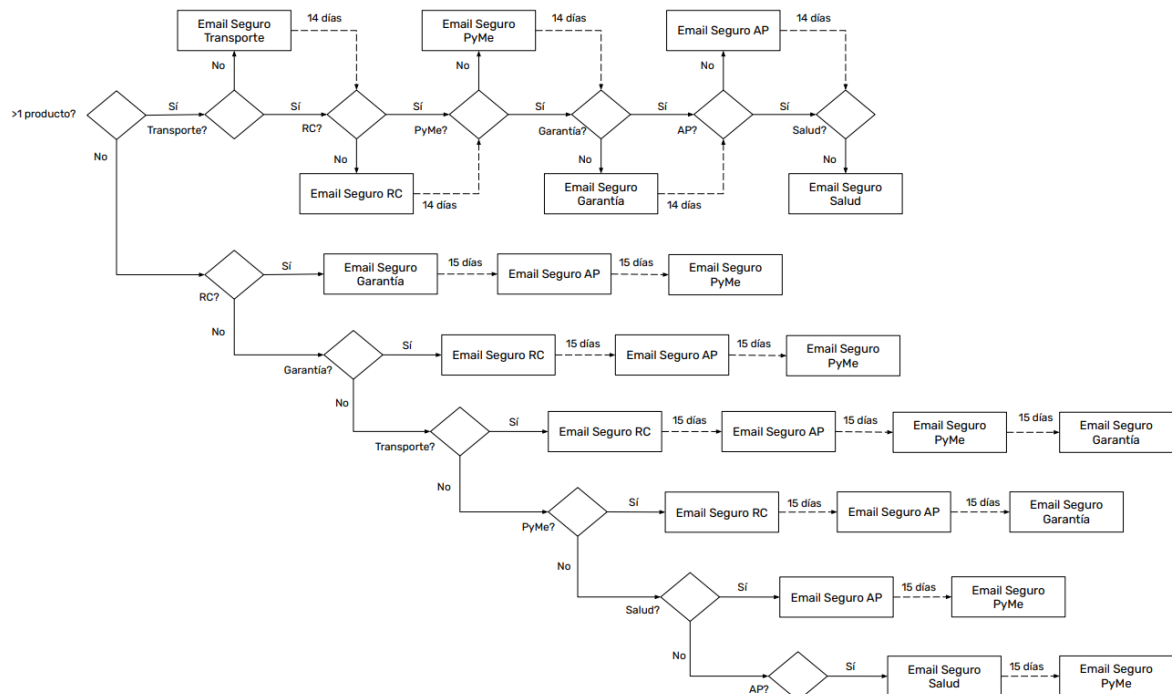


Figura 12: “Diagrama de flujo de Cross-Selling para segmento de mediano valor potencial”

Fuente figura 12: “Creación propia a partir de patrones, datos obtenidos del árbol de decisiones elaborado anteriormente y ayuda del equipo de Marketing”

