



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

**DISEÑO DE UN PROCESO DE VISITAS PROACTIVAS A CLIENTES CON PROBLEMAS EN  
EL SERVICIO DE INTERNET FIJO**

PROYECTO DE GRADO PARA OPTAR AL GRADO DE MAGÍSTER EN INGENIERÍA DE  
NEGOCIOS CON TECNOLOGÍAS DE INFORMACIÓN

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

**JOSÉ VICTOR TERRAZAS TORRICO**

PROFESOR GUÍA:  
MARCEL GOIC FIGUEROA  
PROFESOR COGUÍA:  
SEBASTIÁN RÍOS PÉREZ  
COMISIÓN:  
CARLA RÍOS PALMA

SANTIAGO DE CHILE  
2021

RESUMEN DE LA MEMORIA PARA OPTAR  
AL TÍTULO DE MAGÍSTER EN CIENCIAS  
DE LA INGENIERÍA  
POR: **JOSÉ VICTOR TERRAZAS TORRICO**  
FECHA: 2021  
PROF. GUÍA: NOMBRE PROFESOR

## **DISEÑO DE UN PROCESO DE VISITAS PROACTIVAS A CLIENTES CON PROBLEMAS EN EL SERVICIO DE INTERNET FIJO**

Debido a la pandemia provocada por el virus SARS-CoV-2 los reclamos del servicio de internet fijo aumentaron de forma inesperada, ya que cambia la relación entre el cliente y el servicio tomando esta una mayor relevancia en la vida del cliente y vislumbrando problemas que antes no afectaban de la misma forma. Este cambio en la necesidad del cliente trae como resultado un aumento de la fuga de clientes de este servicio afectando financieramente a la organización.

Para solucionar el problema se propone el diseño de un proceso de visitas técnicas proactivas a clientes que tengan problemas con el servicio de internet fijo, cambiando la forma de atención y la relación que se tiene con el cliente. Esta solución se basa en detectar que clientes tienen problemas de internet y contactar a dichos clientes de forma proactiva para solucionarlos evitando el reclamo posterior y previniendo un alza inesperada de clientes que tengan problemas y pidan una visita técnica.

La solución apunta a mejorar el servicio de internet y mostrar una preocupación de la organización al cliente, disminuyendo los reclamos de estos y reduciendo la posterior intención de cambiar de proveedor del servicio.

Con este proceso es posible cambiar la forma de atención de los clientes realizando reparaciones en los domicilios antes de que el cliente las detecte y pudiendo balancear la dotación de técnicos en el tiempo previniendo también un alza inesperada en la demanda de estos.

# Agradecimientos

Al llegar al final de un largo proceso como este, no solo del trabajo de título, sino que de todos los años de estudio que conlleva la carrera. Uno se pone a pensar lo difícil que sería sin el cariño, apoyo y compañía de muchas personas que hicieron de esta un gran experiencia y un gran aprendizaje.

En primer lugar quiero comenzar agradeciendo a mis padres por darme la posibilidad de estudiar y por su apoyo constante e incondicional. A mi hermano por su compañía y por motivarme. A mi abuelita Menche por toda su guía, su cariño y amor.

A Gabriela por acompañarme sobretodo en el final de la carrera e impulsarme a trabajar en la tesis cuando la motivación decae. Gracias por tu infinito amor, comprensión y risas.

También es necesario agradecer a los amigos y amigas que hice en mi paso por la universidad, por todos los momentos de felicidad vividos, por esas juntas de estudio, las tardes después de clases, las conversaciones infinitas sobre mil temas y por hacer de la universidad una gran e inolvidable experiencia. Espero que la vida siga reuniéndonos y compartiendo muchos momentos.

A mis amigos y migas de la vida, que también siguento acompañándome durante todos los años, a pesar de que no nos veíamos con la misma frecuencia que en el colegio siempre permanecieron a mi lado y me brindaron su apoyo cuando lo necesitaba. Espero que la vida siga reuniéndonos.

Quiero agradecer a las y los docentes que me tuve el agrado de conocer durante mi paso por la universidad, tanto en pregrado, como en postgrado.

También quiero agradecer a todas las personas que he tenido el gusto de compartir en los equipos de trabajo, comenzando por las personas que se convirtieron en una pequeña familia para mi en Penta Analytics, la cual fue mi primer trabajo. Personas que me dieron sus consejos y su guía. También agradecer a los equipos de trabajo que tuve y que tengo en el trabajo actual, donde también tuve la oportunidad de realizar el presente trabajo de título. Gracias por la oportunidad y el aprendizaje que me han brindado.

Agradecer a mi profesor guía Marcel por facilitar a que este trabajo llegara a buen puerto y también a la señora Lisien por ayudarme con la redacción del trabajo y sus consejos.

Finalmente, gracias por todas las oportunidades que he tenido y por todo el apoyo que me han brindado. Espero poder ejercer mi profesión de la mejor manera y siempre teniendo en cuenta el rol social de ésta.

# Tabla de Contenido

<b>1. Introducción y contexto</b>	<b>1</b>
1.1. Antecedentes de la industria . . . . .	2
1.2. Descripción general de la empresa . . . . .	3
1.3. Acerca del problema y su justificación . . . . .	4
1.4. Objetivo general y resultados esperados del proyecto . . . . .	6
1.4.1. Objetivo general . . . . .	7
1.4.2. Objetivos Específicos . . . . .	7
1.4.3. Resultados Esperados . . . . .	7
1.4.4. Alcance . . . . .	8
1.4.5. Riesgos potenciales . . . . .	8
<b>2. Marco Teórico</b>	<b>11</b>
2.1. Metodología del proyecto . . . . .	11
2.2. Marco Teórico para la Lógica de Negocios . . . . .	12
2.2.1. CRISP-DM . . . . .	12
2.2.2. Machine learning . . . . .	14
2.2.3. Algoritmos supervisados: . . . . .	15
2.2.3.1. Regresión Logística . . . . .	15
2.2.3.2. Árboles de decisión . . . . .	15
2.2.3.3. Máquinas de soporte vectorial (SVM) . . . . .	15
2.2.3.4. Random forest . . . . .	16
2.2.3.5. XGBoost . . . . .	17
2.2.4. Métricas de desempeño . . . . .	17
2.2.4.1. Matriz de confusión . . . . .	17
2.2.4.2. Área bajo la curva ROC . . . . .	19
2.2.4.3. Prueba de Kolgomorov-Smirnov . . . . .	19
<b>3. Planteamiento estratégico y análisis de la situación actual</b>	<b>21</b>
3.1. Posicionamiento estratégico . . . . .	21
3.2. Modelo de negocios . . . . .	22
3.3. Diagnóstico de la Situación Actual . . . . .	24
3.3.1. Problema Identificado / Oportunidad identificada . . . . .	25

3.3.2. Arquitectura de Procesos AS IS . . . . .	28
3.3.3. Modelamiento Detallado de Procesos AS IS . . . . .	29
3.4. Cuantificación del Problema . . . . .	30
<b>4. Propuesta de rediseño de procesos</b>	<b>32</b>
4.1. Direcciones de cambio y alcance . . . . .	32
4.2. Propuesta de solución . . . . .	33
4.2.1. Modelamiento Detallado de Procesos TO BE . . . . .	34
4.2.2. Diseño de Lógica de Negocios . . . . .	35
4.3. Diseño del modelo de detección de problemas del servicio de internet fijo . . . . .	37
4.3.1. Entendimiento de los datos . . . . .	38
4.3.2. Preparación de los datos . . . . .	42
4.3.3. Modelamiento . . . . .	43
4.4. Evaluación de los modelos . . . . .	44
4.5. Resultados del modelo seleccionado . . . . .	45
<b>5. Propuesta de Apoyo Tecnológico</b>	<b>47</b>
<b>6. Evaluación del Proyecto</b>	<b>51</b>
6.1. Evaluación Técnica . . . . .	51
6.2. Evaluación Económica . . . . .	54
6.2.0.1. Análisis de sensibilidad . . . . .	56
<b>7. Conclusión y Trabajos Futuros</b>	<b>58</b>
<b>Bibliografía</b>	<b>60</b>
<b>Anexo A. Gráficos univariados de métricas vs tasa de llamados</b>	<b>62</b>
<b>Anexo B. Resultados del modelo</b>	<b>66</b>

# Índice de Tablas

1.1.	Informe trimestral del sector de telecomunicaciones . . . . .	2
1.2.	Variación en participación de mercado - internet fijo . . . . .	3
3.1.	Cuantificación del problema . . . . .	31
4.1.	Variables de diseño . . . . .	33
4.2.	Métricas de desempeño de los cable módem . . . . .	38
4.3.	Llamadas . . . . .	38
4.4.	Cartera de clientes . . . . .	38
4.5.	Umbral operaciones vs umbrales propuestos en las métricas de desempeño de los cable módem . . . . .	42
4.6.	Ejemplo selección de variables índice ks . . . . .	43
4.7.	Evaluación algoritmos de clasificación . . . . .	44
6.1.	Evaluación técnica - mirada retrospectiva . . . . .	52
6.2.	Evaluación técnica - resultados . . . . .	52
6.3.	Evaluación técnica - mirada retrospectiva 2 . . . . .	53
6.4.	Evaluación técnica - resultados 2 . . . . .	54
6.5.	Evaluación técnica - resultados grupo en buen estado . . . . .	54
6.6.	Parámetros para la evaluación económica . . . . .	55
6.7.	Flujo de caja mensual . . . . .	56
6.8.	Flujo de caja anual . . . . .	56
6.9.	Análisis de sensibilidad . . . . .	57

# Índice de Ilustraciones

1.1.	Evolución penetración internet [Subtel, 2019]. . . . .	3
1.2.	Evolución tráfico internet fijo [Subtel, 2020]. . . . .	5
1.3.	Matriz de riesgos. . . . .	10
2.1.	Modelo CRISP-DM [Wirth and Hipp, 2000]. . . . .	13
2.2.	Support Vector Machine. [Mohri et al., 2018] . . . . .	16
2.3.	Random Forest. . . . .	17
2.4.	Matriz de confusión. . . . .	18
3.1.	Modelo Delta de Hax. . . . .	21
3.2.	Modelo de negocios de VTR. . . . .	23
3.3.	Evolución de la tasa de llamadas por motivos técnicos. . . . .	25
3.4.	Evolución de la tasa de fuga del servicio de internet. . . . .	25
3.5.	Árbol Why - Aumento desconexiones. . . . .	26
3.6.	Árbol How - Mejorar calidad de servicio. . . . .	27
3.7.	Venta y atención de clientes de internet fijo. . . . .	29
3.8.	Proceso de resolución de incidentes técnicos. . . . .	30
4.1.	Propuesta proceso de visitas domiciliarias proactivas. . . . .	35
4.2.	Ejemplo del supuesto de llamadas técnicas en la detección de clientes . . . . .	37
4.3.	Comparación llamadas vs llamadas técnicas de internet. . . . .	39
4.4.	CER UP Max vs Llamadas técnicas. . . . .	40
4.5.	T3 Timeouts vs Llamadas técnicas. . . . .	40
4.6.	Tasa de llamados vs RX Power en Upstream . . . . .	41
4.7.	Precisión del modelo vs puntos de corte . . . . .	45
4.8.	Importancia de las variables . . . . .	46
5.1.	Propuesta de arquitectura TI del proyecto . . . . .	47
5.2.	Contacto proactivo por mensaje de whatsapp . . . . .	49
5.3.	Plataforma BEC - Ingresar rut del cliente . . . . .	49
5.4.	Plataforma BEC - Campaña proactiva cargada . . . . .	50
6.1.	Cantidad de TR realizados mensualmente . . . . .	53
A.1.	Tasa de llamados vs CCER en Downstream . . . . .	62
A.2.	Tasa de llamados vs CCER en Upstream . . . . .	62
A.3.	Tasa de llamados vs CER en Downstream . . . . .	63
A.4.	Tasa de llamados vs CER en Upstream . . . . .	63

A.5.	Tasa de llamados vs SNR en Downstream	63
A.6.	Tasa de llamados vs SNR en Upstream	64
A.7.	Tasa de llamados vs T3 Timeouts	64
A.8.	Tasa de llamados vs T4 Timeouts	64
A.9.	Tasa de llamados vs RX Power en Downstream	65
A.10.	Tasa de llamados vs RX Power en Upstream	65
A.11.	Tasa de llamados vs TX Power en Upstream	65
B.1.	Curva de ganancia del modelo seleccionado	66



# Capítulo 1

## Introducción y contexto

El trabajo que se expone a continuación está situado en la industria de las telecomunicaciones en Chile, específicamente enfocado en el servicio de internet fijo, lo que tiene gran impacto en las vidas cotidianas y en el avance de la digitalización a nivel mundial y nacional.

Este trabajo expone una problemática que tiene su origen en la pandemia provocada por el virus SARS-CoV-2 conocido como coronavirus. Debido al virus, comienza en todo el mundo un proceso de cuarentenas obligatorias impuestas por los gobiernos. Este hecho genera un cambio en la forma de estudiar, trabajar, reunirse con las personas, entre otras actividades, que pasan desde un plano presencial a un plano digital. Este fenómeno lleva al límite a todas las compañías de telecomunicaciones en Chile y el mundo, debido a que no estaban preparadas para tan repentino aumento en el volumen de tráfico. Esto es solo la primera parte del problema, ya que expone a las compañías no solo a problemas de capacidades, sino también a problemas en la calidad del servicio, tomando éste una prioridad que antes no tenía debido al aumento de las exigencias del cliente.

En este contexto, la organización VTR Comunicaciones SpA, de ahora en adelante VTR, es una de las principales afectadas con los efectos de la pandemia, por cuanto han aumentado los reclamos y la fuga de los clientes del servicio de internet fijo. Como respuesta a este problema de negocio se propone el diseño e implementación de un proceso de acciones proactivas para mantener la calidad del servicio de internet fijo en niveles que no dañen la experiencia de los clientes.

El trabajo está basado en la metodología de ingeniería de negocios [Barros, 2015] que permite vincular la estrategia de la empresa y los procesos de la misma. También se utiliza la metodología de CRISP-DM [Wirth and Hipp, 2000] para realizar el diseño y desarrollo de un modelo que permita realizar las acciones de forma proactiva.

## 1.1. Antecedentes de la industria

La industria de telecomunicaciones en Chile está regulada por la subsecretaría de Telecomunicaciones, organismo dependiente del ministerio de Transporte y Telecomunicaciones.

Los servicios que comprende la industria de telecomunicaciones en Chile son internet fijo, internet móvil, telefonía fija, telefonía móvil, televisión de pago y larga distancia internacional. En la Tabla 1.1 se observa la variación interanual entre el primer trimestre del 2019 y el primer trimestre del 2020 para cada uno de los servicios antes mencionados. En esta tabla se observa un aumento de 3,8% en internet fija y un aumento de 2,2% en internet móvil (3G + 4G) y 7,5% en 4G siendo estos los únicos servicios donde observa un aumento entre ambos trimestres. Este dato le entrega relevancia al servicio de internet fija debido al crecimiento que presenta en el último periodo, siendo éste el foco del proyecto descrito en el presente informe.

Tabla 1.1: Informe trimestral del sector de telecomunicaciones

<b>Servicio</b>	<b>Unidad</b>	<b>1T - 2020</b>	<b>Variación interanual</b>
Internet Fija	Miles	3.489	3,8 %
Internet Móvil (3G + 4G)	Miles	18.946	2,2 %
4G	Miles	16.600	7,5 %
Telefonía Fija	Miles	2.707	-7,5 %
Telefonía Móvil	Miles	24.772	-3,7 %
TV Pago	Miles	3.221	-2,8 %
Larga Distancia Internacional	Millones de minutos	5,2	-27 %

Para entender la evolución del servicio de internet fijo y en general el servicio de internet. Se puede observar la Figura 1.1. En dicha figura se muestra los miles de accesos al servicio de internet cada 100 habitantes, donde se observa que hay 116,1 conexiones a internet fijo y móvil por cada 100 habitantes al cierre del año 2019 [Subtel, 2019], es decir, más de una conexión por habitante.

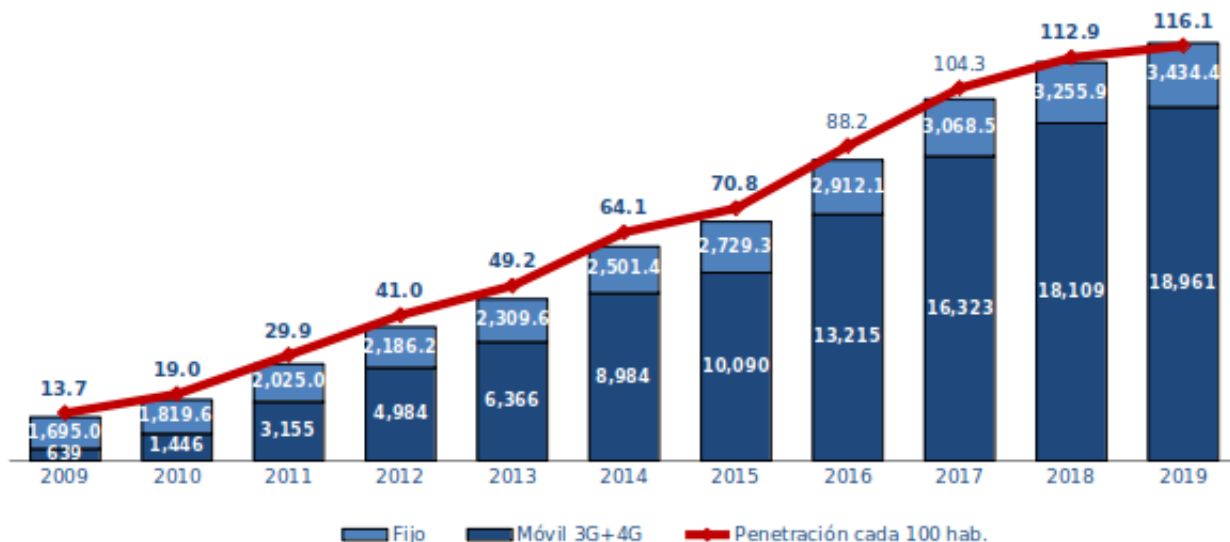


Figura 1.1: Evolución penetración internet [Subtel, 2019].

Otro factor importante para entender la industria de telecomunicaciones corresponde a los diferentes actores que compiten en este mercado. Dentro del servicio de internet fijo el 99,6 % del mercado se reparte entre Movistar, VTR, Claro, Grupo GTD, Mundo pacífico y grupo Entel, como se puede observar en la Tabla 1.2. Agregando a la competencia desde el sector de internet móvil se encuentra WOM con un 23,6 % de participación de mercado, el que actualmente compite con su oferta de internet móvil para hogares.

Tabla 1.2: Variación en participación de mercado - internet fijo

Empresa	mar-19	mar-20
<b>Movistar</b>	29,9 %	26,5 %
<b>VTR</b>	37,9 %	38,2 %
<b>Claro</b>	13,9 %	13,2 %
<b>GTD</b>	7,9 %	7,9 %
<b>Mundo Pacífico</b>	5,6 %	7,6 %
<b>ENTEL</b>	4,2 %	6,2 %
<b>Otros</b>	0,6 %	0,4 %

## 1.2. Descripción general de la empresa

VTR es una empresa que provee servicios de telecomunicaciones, tales como acceso a internet, telefonía y televisión digital, para clientes individuales y empresas pequeñas y medianas.

Desde el año 2018 VTR pertenece en un 100 % a Liberty Latin America, empresa

estadounidense líder en telecomunicaciones, con presencia en diferentes países de Latinoamérica y el Caribe tales como Chile, Panamá, Costa Rica, entre otros.

La misión de VTR es “**Conectar comunidades y transformar vida**”. Esta misión muestra la importancia de la experiencia del cliente en la compañía, poniendo el foco en el cliente. También se puede observar esta característica en la visión de VTR que corresponde a “**Innovar para crear momentos significativos para los clientes, generando crecimiento en el mercado, con una visión, una cultura y un equipo**”[VTR, 2018].

En Chile, VTR es la compañía líder en servicios de internet fijo y televisión digital. Servicios en los que logra una participación en el mercado de 38,2% y 24%, respectivamente, durante el primer trimestre del 2020; VTR también corresponde a la segunda compañía con más participación de mercado en telefonía fija con una participación del 20,5%, siendo superada únicamente por Telefónica. También participa en el mercado de telefonía móvil e internet móvil con un pequeño porcentaje del mercado de un 1,2% y un 1,5% respectivamente.

### **1.3. Acerca del problema y su justificación**

La crisis sanitaria provocada por el virus SARS CoV-2, ha generado en todo el mundo un nuevo estilo de vida. Debido a las características de transmisión del virus y las cuarentenas impuestas por los gobiernos, ha cambiado para la mayoría de la población la forma de estudiar, de trabajar, de reunirse, entre otras actividades, provocando así un aumento en el tráfico de datos de internet fijo. Entre febrero y abril del 2020 se observa un 40% de aumento en promedio [Subtel, 2020], aumento que supera incluso el crecimiento en tráfico de 1 año en situación normal, según datos anteriores. Esto se puede observar en la Figura 1.2, entre enero del 2019 y enero del 2018 existe un aumento de solo el 17%.

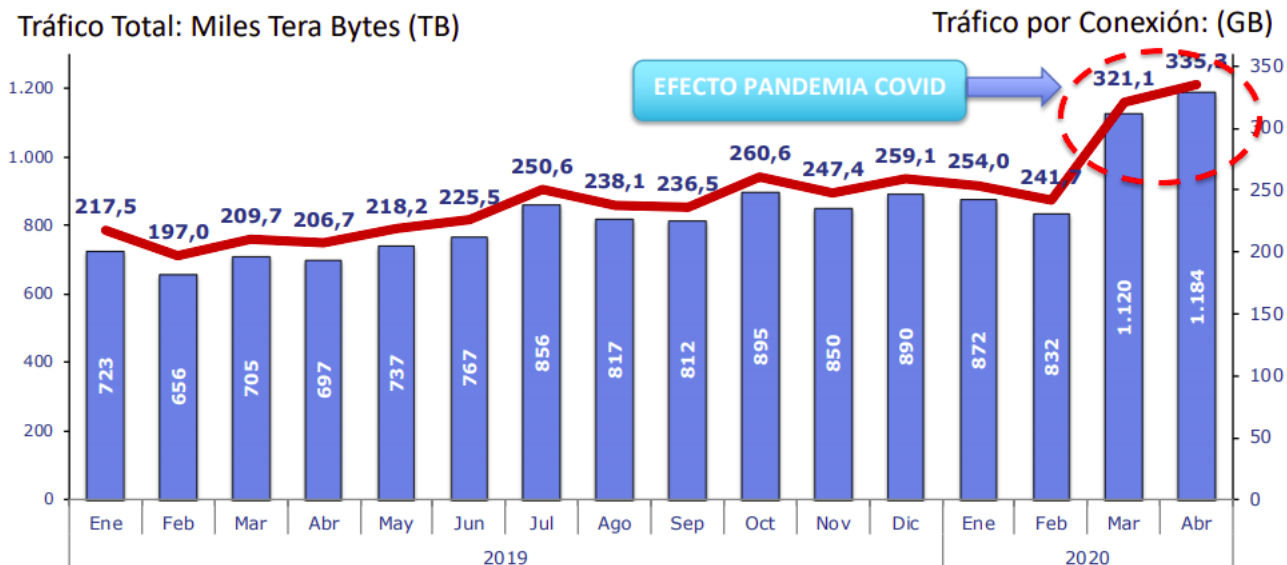


Figura 1.2: Evolución tráfico internet fijo [Subtel, 2020].

El crecimiento repentino en el tráfico de internet ha tenido como consecuencia una baja en la calidad del servicio de internet entregado a los clientes, debido a que las carreteras que soportan el tráfico de internet no están adecuadas para ese nivel de tráfico y ese aumento inesperado. Este malestar con la experiencia del cliente ha sido causado no solo por problemas de capacidades, sino también, porque se han puesto en evidencia fallas en la calidad del servicio al tener un usuario más estricto y con un uso más recurrente, lo que se refleja en los 9.605 reclamos ingresados entre el 18 de marzo y 17 de abril del 2020 al Departamento de Gestión de Reclamos (DGR) de Subtel; superando con creces los 5.396 recibidos durante el mismo periodo del 2019. Una de las principales empresas golpeadas por este fenómeno es VTR, la cual tiene un aumento del 100 % de llamadas por motivos técnicos en los primeros meses de la pandemia y un 33 % más de clientes que dan de baja el servicio de internet en el mes de junio, seguido de un 66 % más en los meses posteriores. Como respuesta, VTR ha desencadenado un plan de ampliación de capacidades que se ha llevado a cabo desde inicios de la cuarentena en el país terminando a fines de abril, resolviendo en gran parte el problema de capacidades.

Sin embargo, pese a que los trabajos han ayudado a descongestionar las redes y que los problemas de capacidades se hayan visto reducidos, las llamadas por motivos técnicos y reclamos no han disminuido como se esperaba. Esto es debido a que las capacidades no son los únicos factores que provocan las fallas en el servicio de internet, también se detectan otros problemas como: cortes de red, robos de cables, falta de proactividad para mantener en un buen estándar los niveles de servicio y un mal desempeño de la red. La red de VTR está compuesta

en un 100 % de conexiones HFC que son una combinación de fibra óptica hasta el nodo y cable coaxial entre el nodo y el cable módem del cliente, esta tecnología requiere mayor mantención que la red de fibra óptica, siendo un factor relevante en los reclamos del cliente.

Al consultar a los clientes de VTR sobre los principales motivos de su fuga se observa que un gran porcentaje es debido a problemas técnicos con el servicio y descontento con la calidad de servicio recibida. Además, la literatura indica que existe una influencia de la calidad del servicio en la intención de cambiar de proveedor de internet [Dwivedi et al., 2010]. Esto implica que gran porcentaje de la fuga de clientes se explica por la mala experiencia de los clientes, abriendo espacio para una mejora en este ámbito.

Actualmente, el servicio técnico por parte de VTR tiene una respuesta reactiva, es decir, el cliente tiene que sentir el malestar en el servicio y luego contactar a VTR para que recién se inicie el proceso de atención y resolución de problemas técnicos. Esto significa que el cliente tiene que sentir la degradación en la experiencia para que esta sea solucionada por VTR.

Una posible solución al problema antes mencionado consiste en generar alertas tempranas de los problemas técnicos que puedan percibir los clientes y a partir de estas alertas gestionar acciones que eviten los reclamos de los clientes y mantenga una mejor calidad del servicio.

## **1.4. Objetivo general y resultados esperados del proyecto**

El proyecto propuesto consiste en diseñar un proceso para contactar a los clientes con problemas en el servicio de internet fijo y/o una degradación en sus métricas de desempeño, de forma temprana, es decir, antes que los clientes inicien un reclamo por el problema y/o incluso antes que el cliente pueda percibir el problema en su domicilio.

La motivación de realizar esta tarea de forma temprana es evitar que el cliente se sienta molesto por la degradación en el servicio, esto se debe principalmente a que debido a la pandemia el uso de internet toma una importancia preponderante en la vida de las personas y las ayuda a realizar las tareas que antes realizaban sin necesidad de esta herramienta, como estudiar, trabajar, jugar, comunicarse con otras personas u otras tareas. Por esto es necesario mantener en niveles óptimos el servicio de internet fijo, para que no se suspendan las clases, no haya interrupciones en medio de una entrevista de trabajo o en medio de una reunión de trabajo, lo que significa una molestia de los clientes con la compañía proveedora del servicio de internet.

Por otro lado, realizar visitas de forma proactiva permite balancear las cargas tanto

del call center, como la de los técnicos que realizan las visitas a domicilio. De esta forma, solucionar problemas en forma distribuida en el tiempo y prevenir algún peak que no soporte la dotación actual.

Para realizar el proyecto es necesario conocer la dotación tanto de los canales de contacto con los clientes, como la dotación de técnicos para visitas domiciliarias. También es necesario diseñar un modelo de detección y priorización de problemas de clientes, el que a partir de las métricas de desempeño recolectadas desde los cable módem de los clientes detecte quienes tienen problemas con el servicio (ya sea que estos se den cuenta o no) y, posteriormente, priorizar según quienes tengan un problema más crítico, para dirigir las acciones proactivas a aquellos clientes con mayor problema. También como medida fidelizadora se propone realizar una certificación total de todos los servicios contratados por el cliente.

#### **1.4.1. Objetivo general**

El objetivo general del proyecto es el diseño de un proceso de visitas proactivas a clientes con problemas en el servicio de internet fijo, para mantener la calidad del servicio en niveles operacionales, es decir, se busca contactar a clientes que se observen con problemas en el servicio de internet antes que el cliente genere el reclamo y/o antes que el cliente se dé cuenta del problema y tenga la degradación en la experiencia del servicio.

#### **1.4.2. Objetivos Específicos**

1. Analizar y conocer la situación actual de las llamadas por motivos técnicos de internet fijo y la resolución de incidentes por parte de VTR.
2. Diseñar y desarrollar un modelo que determine que clientes tienen problemas con el servicio de internet fijo.
3. Diseñar el flujo de las actividades del proceso de visitas proactivas a clientes con problemas en el servicio. Desde como detectar a los clientes, como elegir a quienes contactar y la forma de contactarlos.
4. Diseñar un modelo que a partir de las métricas de desempeño de los módem de los clientes pueda detectar y priorizar aquellos que tienen problemas con el servicio de internet fijo.
5. Evaluar la rentabilidad del proyecto presentado a partir de una análisis económico.

#### **1.4.3. Resultados Esperados**

Al término de este proyecto se espera el cumplimiento de los siguientes resultados:

1. Comprensión de la situación actual y los requerimientos del proceso como la dotación del call center, dotación de técnicos, etc.
2. Descripción del flujo de actividades a partir de un BPMN, conociendo las tareas y los responsables de cada una de estas.
3. El diseño y desarrollo de un modelo de priorización de clientes con problemas en el servicio de internet fijo que ordene desde aquellos con problemas más críticos a los que no tienen problemas.
4. Determinación de los elementos de la evaluación económica del proyecto con la sensibilidad de las variables más importantes.

#### **1.4.4. Alcance**

El proyecto está situado en la cadena de valor de VTR, específicamente en la administración y relación con los clientes. Dentro de este proceso se sitúa en la venta y atención de clientes de internet y en un nivel más abajo en el servicio al cliente de internet fijo.

Los problemas que se van a detectar son únicamente a partir de las métricas de desempeño de los cable módem disponibles gracias a la plataforma Servassure que realiza estas mediciones cada hora. No se van a considerar problemas de capacidades, ni problemas puntuales conocidos como masivos dentro de los cuales están los cortes de luz, cortes de fibra óptica, robo de cables y caídas masivas.

El modelo desarrollado durante el proyecto se encarga únicamente de identificar los clientes objetivos para el proyecto de visitas proactivas.

Y finalmente, el proyecto no incluye el diseño y ejecución de las visitas técnicas domiciliarias, estas serán realizadas de forma tradicional con el único plus de certificar todos los servicios contratados por el cliente, no solamente el servicio de internet, es decir, que para cerrar la visita el técnico tiene que certificar que todos los servicios del cliente funcionan de forma correcta.

#### **1.4.5. Riesgos potenciales**

Los potenciales riesgos se describen a continuación:

1. La no aceptación del cliente a la visita domiciliaria: es un riesgo importante, porque VTR se va a proponer una visita técnica a aquellos clientes que se vean con problema y que no hayan generado algún reclamo, por esto es posible que el cliente no quiera aceptar la visita técnica. Otro motivo para no aceptar la visita técnica es la pandemia del Covid-19, ya que pueden haber personas de riesgo en el hogar. Para este caso se va a tratar con todas las



medidas sanitarias necesarias para no poner en riesgo las personas del hogar y si de todas formas el cliente no quiere aceptar la visita por que no la considera necesaria, este será marcado en la plataforma de atención a clientes en caso que se contacte con VTR por problemas técnicos tener prioridad y un trámite ágil para agendar la visita técnica.

2. La dotación de técnicos: es un riesgo importante la dotación necesaria de técnicos para realizar las visitas domiciliarias, esto es importante ya que por un lado si se agendan más visitas de las que se pueden realizar va a afectar de forma negativa en la percepción del servicio por parte del cliente, ya que le proponemos una visita que no es posible realizar. Para esto es necesario conocer la tasa de aceptación de visitas y la capacidad de visitas disponibles, para una mayor flexibilidad ir reevaluando las cantidades de forma semanal o mensual para llegar a una meta y no caer en el problema antes citado.
3. La detección de problemas conlleva varios desafíos y riesgos: el primero corresponde a la falta de información respecto a cuándo hay un problema y cuándo no, es decir, identificar problemas de internet es algo difuso y no se tiene una etiqueta que lo indique; esto se resuelve por aproximación, considerando el llamado del cliente a la plataforma técnica de internet. Esto bajo el supuesto que cuando un cliente tiene problemas con su servicio de internet llama a la plataforma técnica de VTR para solucionarlo y poner un reclamo generando un ticket. Entonces, este ticket se utiliza para aproximar problemas en el servicio de internet. Otro desafío consiste en definir bajo qué umbral se determina si un cliente tiene problemas con el servicio o no; el riesgo que esto conlleva es que si el umbral es muy estricto, solo un pequeño grupo de clientes se identificarán con problemas y habrán clientes con problemas que no serán etiquetados como tal. Por el contrario, si el umbral es muy laxo, una gran cantidad de clientes tendrán problemas y será difícil priorizar o determinar acciones para grupos tan grandes de clientes.
4. El Ruido: otro riesgo es el ruido que existe en los datos, dado que se utiliza las llamadas como una aproximación a los problemas del cliente, y el caso es que no todos los clientes llaman cuando tienen problemas y que no todos los que llaman tienen problemas realmente. Para reducir este riesgo se filtrarán las llamadas, considerando solo las llamadas que entren por la plataforma técnica. También, existe el ruido en los datos extraídos de los cable módem, ya que estos son perceptibles a ruido y variaciones anormales, para ello se realizarán transformaciones que hagan más robusta a la data.
5. La implementación: para ser implementado de buena forma tiene que haber apoyo de cada una de las áreas involucradas y estas tienen que ser parte del diseño del proceso, lo cual tiene grandes implicancias ya que si el proyecto no es apoyado por alguna de las áreas relevantes se estanca y no avanza.

Las áreas relevantes en el proceso son: el área de BI que tiene que realizar los desarrollos para mostrar a los clientes que van a ser priorizados con problemas, el área de clientes que se encarga del contacto con los clientes y de los canales de atención disponibles, como también de la dotación necesaria para atender las consultas de los clientes, y finalmente el área técnica que administra la visita técnica domiciliaria y se encarga de la dotación disponible para las visitas y su ejecución.

Probabilidad/Impacto	Insignificante	Menor	Critica	Mayor	Catastrófico
Constante		3-4			
Moderado		1		2-5	
Ocasional					
Posible					
Imposible					

Figura 1.3: Matriz de riesgos.

En la Figura 1.3 se ubican los riesgos antes citados de acuerdo a su probabilidad de ocurrencia y al impacto que tienen en el negocio. Según se observa, los riesgos más críticos del proyecto son referidos al personal, el 2 y el 5. El riesgo número 5 corresponde a la falta de apoyo de las áreas involucradas lo que atrasaría el proyecto y dificultaría su implementación, para disminuir este riesgo es importante involucrar a las áreas que participan en el diseño del proyecto y que de esta manera tengan propiedad sobre el mismo y lo impulsen. El riesgo más crítico es el número 2, ya que al no tener la dotación de técnicos para realizar las visitas proactivas, el proyecto no es factible de realizar y agendar el tipo de acciones necesarias y, sin poder cumplirlas, afectaría de peor forma la imagen de la organización.

# Capítulo 2

## Marco Teórico

### 2.1. Metodología del proyecto

La metodología que sigue la realización del proyecto es la ingeniería de negocios [Barros, 2015]. Esta metodología entrega fundamentos y buenas prácticas para diseñar una empresa, describiendo la arquitectura empresarial y los procesos necesarios para la supervivencia y competitividad de la empresa. Esta corresponde a una metodología de tipo “top-down”, lo que permite una visión sistémica de la empresa y de sus procesos, como también permite un enlazamiento coherente desde la estrategia de la empresa hasta la arquitectura empresarial y los procesos más operacionales [Barros, 2015].

Las etapas de esta metodología se muestran a continuación:

#### 1. Planteamiento estratégico:

La primera etapa de esta metodología consiste en entender y definir el planteamiento estratégico de la compañía, ya que esta define la forma de competir y de obtener ventajas competitivas para la empresa. Algunas herramientas útiles son: el Modelo Delta de Arnoldo Hax, las fuerzas de Porter, el BalanceScoreCard, entre otros. La idea es entender la forma de competir de la empresa y dónde hay que poner foco, para que de esta manera se vayan alineando los diseños a la estrategia.

#### 2. Definición del modelo de negocio:

El modelo de negocio ayuda a materializar las ideas descritas en la parte anterior, también explica en forma resumida quiénes son los clientes de la empresa, como se genera valor de cara a ellos y de qué forma se monetiza este valor.

#### 3. Diseño de la arquitectura de procesos:

Consiste en instanciar los macroprocesos relevantes de una empresa y determinar las relaciones entre ellas. La idea es que los macroprocesos importantes y su instanciamiento responda al planteamiento estratégico de la

empresa y a su modelo de negocio, asegurando de esta manera la creación de valor a partir de los procesos y la generación de ventajas competitivas.

4. **Diseño detallado de procesos de negocios:** El diseño de los procesos de negocios está determinado por la arquitectura de procesos definida en la parte anterior, donde cada uno de los macroprocesos determinados anteriormente tiene entradas y salidas definidas. Entonces, para la estructura detallada de los procesos se utilizan los patrones de procesos (PPN) que se ahondará más adelante.
5. **Diseño de las aplicaciones TI:** De acuerdo a los procesos de negocios que se detallan en la etapa anterior, ahora se presentan los requerimientos de las herramientas TI que apoyarán los procesos. Así se define en los modelos de procesos la mantención de estado.
6. **Construcción e implementación:** En esta etapa se llevan a la práctica los procesos y aplicaciones TI definidos durante las etapas anteriores. Cabe destacar la importancia de la gestión del cambio para que la implementación funcione de buena forma, ya que tienen que generarse las condiciones para que el cambio resulte. La gestión del cambio no comienza en esta etapa, ya que los mismos diseños se van acordando con los usuarios en los puntos anteriores.

## 2.2. Marco Teórico para la Lógica de Negocios

Dentro de la metodología de ingeniería de negocios [Barros, 2015] en la etapa de diseño de las aplicaciones TI, es necesario presentar las herramientas tecnológicas que apoyarán al proyecto, siendo la principal herramienta de apoyo un modelo de detección y priorización de clientes con problemas en el servicio de internet fijo el cual está dentro del marco de un proyecto de datos. Para realizar este proyecto de datos en el marco del proyecto principal, se utilizará una metodología comúnmente usada para proyectos de data science como es CRISP-DM [Wirth and Hipp, 2000].

### 2.2.1. CRISP-DM

CRISP-DM es una framework desarrollado en 1999 por un grupo de empresas europeas a partir de diferentes versiones de KDD. El objetivo de este modelo de referencia es guiar con las buenas prácticas en proyecto de data science [Wirth and Hipp, 2000].

El modelo de CRISP-DM define una serie pasos y las relaciones entre ellos desde el entendimiento del negocio hasta el despliegue de la solución obtenida. Una

característica destacable es la forma iterativa y circular del modelo, ya que se basa en que uno se puede ir devolviendo a los pasos anteriores, si es necesario.

En la Figura 2.1 se observa el flujo del modelo y a continuación se detallan los pasos a realizar:

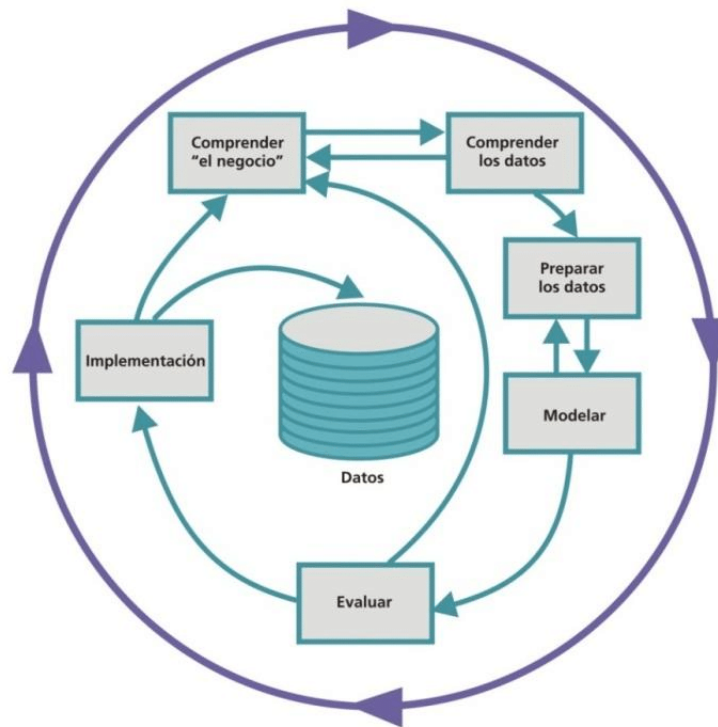


Figura 2.1: Modelo CRISP-DM [Wirth and Hipp, 2000].

- **Comprender el negocio:** La primera etapa tiene como foco entender y definir los objetivos y requerimientos desde la perspectiva del negocio. En esta etapa se describe la situación actual, el problema de negocio y se define qué objetivos deben cumplirse para indicar que se resuelve dicho problema.
- **Comprender los datos:** Esta etapa tiene como objetivo disponibilizar los datos con los que se va a trabajar, garantizar la calidad de los datos y entregar insights que sirvan para describir el problema a partir de los datos y ayuden a resolver el problema de negocio. Cabe destacar que uno se puede devolver al paso anterior si es necesario, por ejemplo, si la calidad de los datos que tengo no cumple con los requisitos mínimos para seguir con el proyecto.
- **Preparar los datos:** Una vez que se conoce el problema de negocio y los datos con los cuales se va a trabajar es necesario procesarlos para el modelamiento. Dentro de esta etapa se considera la selección de los datos con los que se va a trabajar, la limpieza de los datos, unificar las bases con las que se trabajará, crear nuevas características y reestructurar la data a la granularidad requerida para el modelamiento.

- **Modelar:** Dependiendo del problema de negocio y los datos que se tenga disponibles, se define qué tipo de modelos se van a utilizar. Entonces, en esta etapa se define las técnicas de modelamiento y los supuestos del modelamiento, se genera el diseño de evaluación, se construye el modelo y se pone en práctica con los datos.
- **Evaluar:** En esta etapa se evalúa los resultados de la etapa anterior basados en el diseño definido anteriormente, luego se revisa el proceso realizado y se definen los próximos pasos, los cuales dependen de los resultados obtenidos, ya que puede ser necesario volver a alguna de las etapas anteriores hasta a definir el problema de negocio nuevamente.
- **Implementación:** La etapa de despliegue consiste en la puesta en producción del modelo, es decir, la integración de éste con los procesos de negocio de la organización. Para realizar esto es necesario hacer un plan de despliegue, indicar la forma de monitorear y mantener el modelo, y finalmente ejecutar el plan.

### 2.2.2. Machine learning

Los algoritmos utilizados para realizar el modelamiento en el framework anteriormente mencionado (CRISP-DM), corresponden a los algoritmos de machine learning. Machine learning o aprendizaje automático es un enfoque algorítmico para modelar a través de la experiencia, es decir, que aprende de forma automática a partir de los datos empíricos que se entreguen [Mohri et al., 2018].

Los algoritmos de machine learning están divididos en tres grandes campos, dependiendo de los datos que se utilizan y el tipo de tarea a realizar. Estos 3 grupos corresponden a los modelos supervisados, modelos no supervisados y modelos de reforzamiento. Para el proyecto realizado en este trabajo se utilizan modelos supervisados que se describen a continuación:

1. **Modelos supervisados:** Los modelos supervisados son aquellos que al entrenar se conoce la variable objetivo o variable respuesta del problema, ya que se supervisa el aprendizaje de dicha tarea. Este aprendizaje es utilizado para predecir y se divide en 2 tipos dependiendo de la naturaleza de la variable objetivo. Los algoritmos para predecir variables objetivo continuas son llamados modelos de regresión y cuando la variable es categórica se llaman modelos de clasificación.

### **2.2.3. Algoritmos supervisados:**

Como se menciona en la parte anterior los modelos supervisados son aquellos que aprenden un comportamiento a partir de ver ejemplos de los datos y ejemplos de las respuestas del comportamiento para encontrar patrones en los datos que ayuden a predecir cierto comportamiento. Para realizar esta tarea hay diversos algoritmos, entre los más conocidos están: la regresión logística, los árboles de decisión, máquinas de soporte vectorial (SVM), redes neuronales y otros algoritmos de ensamblaje que se construyen en base a estos otros como random forest, xgboost, entre otros. [Murphy, 2012].

A continuación se detallarán algunos algoritmos que serán utilizados:

#### **2.2.3.1. Regresión Logística**

Modelo de regresión logística es un modelo generativo de probabilidades que se utiliza para tareas de clasificación. En este algoritmo se busca responder a la pregunta de cuál es la probabilidad de que una muestra pertenezca a una categoría dado ciertas características. Para este modelo, se utiliza una función sigmoide en el caso de la clasificación binaria y para la generalización multiclase, la función que se utiliza es softmax [Svensén and Bishop, 2007].

#### **2.2.3.2. Árboles de decisión**

Árboles de decisión son algoritmos definidos por particiones recursivas en el espacio descrito por los datos, de esta manera se define regiones para los datos de entrada que separen de forma más homogénea la variable de respuesta. Estas reglas son representadas por un árbol, donde las hojas corresponden a la regiones del espacio [Mohri et al., 2018].

Estos algoritmos se caracterizan por ser algoritmos avaros, es decir, buscan óptimos locales para cada iteración. También, son conocidos por su simpleza al explicar y su alta interpretabilidad, ya que las reglas de división se ven como reglas que hay que considerar para llegar a los grupos finales que son las hojas de los árboles.

#### **2.2.3.3. Máquinas de soporte vectorial (SVM)**

Las máquinas de soporte vectorial son técnicas de aprendizaje supervisado, que en palabras simples, funcionan a partir de generar un hiperplano que logra separar de mejor forma los datos respecto a la variable objetivo. Esta técnica en un

comienzo se diseña para realizar tareas de clasificación binarias, pero hay modificaciones que generan clasificación para múltiples clases y también tareas de regresión [Mohri et al., 2018].

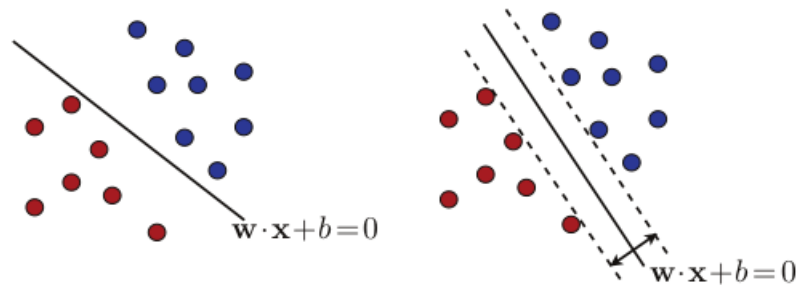


Figura 2.2: Support Vector Machine. [Mohri et al., 2018]

Como se observa en la Figura 2.2, se definen los vectores de soporte a partir de los puntos más alejados de cada categoría y en relación a estos se define a continuación el hiperplano entre medio de ambos vectores de soporte. Entonces, la forma de generar el hiperplano es a partir de maximizar el margen entre los vectores de soporte y minimizar el error de clasificación. En términos matemáticos la función objetivo del problema es [Suykens and Vandewalle, 1999]:

$$\min_{w, \xi_k} \frac{1}{2} w^T w + c \sum_{k=1}^n \xi_k \quad (2.1)$$

#### 2.2.3.4. Random forest

Este es un algoritmo de ensamblaje de bagging o agregación basado en árboles de decisión individuales. La idea de este tipo de algoritmos es promediar el ruido de muchos algoritmos aproximadamente insesgados y de esta manera reducir la varianza. En este algoritmo se crea varios árboles de decisión idénticamente distribuido, donde el valor esperado del sesgo del algoritmo de bagging es igual al valor esperado del sesgo de cada árbol de decisión de forma individual, por lo que la ganancia viene por el lado de reducir la varianza y evitando que un árbol se sobreajuste únicamente a los datos entregados y no pueda generalizar a nuevas muestras [Hastie et al., 2009].

En la Figura 2.3 se muestra un esquema de este algoritmo, en el cual a partir de partes del dataset se van generando múltiples árboles de decisión y luego se combinan las salidas de cada algoritmo teniendo el resultado final.



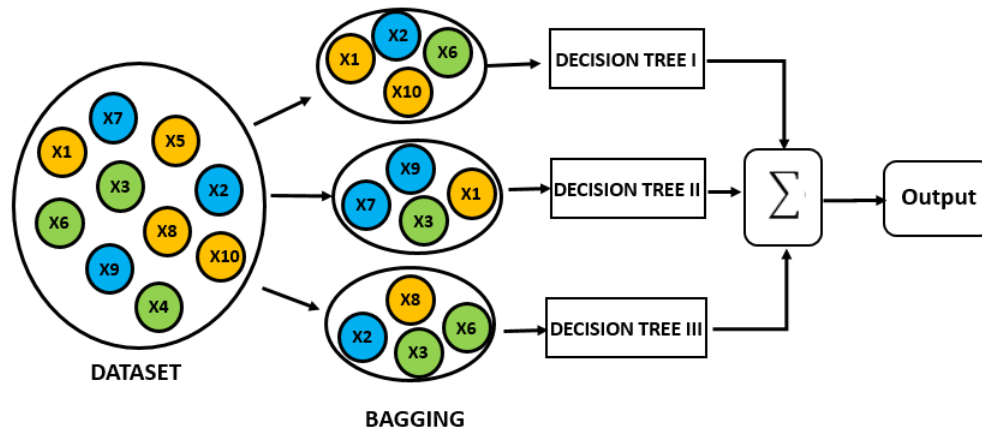


Figura 2.3: Random Forest.

### 2.2.3.5. XGBoost

Xgboost corresponde a un algoritmo de ensamblaje secuencial o de boosting basado en árboles de decisión individuales. A diferencia del random forest este algoritmo no combina la salida de múltiples árboles de decisión, sino que realiza mejoras en forma secuencial y cada árbol toma la salida del árbol anterior y la mejora a partir de una función de costo y de un algoritmo para optimizar esta función coste que corresponde al descenso de gradiente [Chen and Guestrin, 2016].

En pocas palabras en cada paso se construye un árbol de decisión y se compara con el árbol de decisión anterior, si este tiene mejores resultados se emplea el nuevo árbol, en caso contrario, se modifica el árbol anterior por medio del descenso de gradiente.

## 2.2.4. Métricas de desempeño

Para la evaluación de los modelos propuestos es necesario utilizar métricas que indiquen el desempeño de dichos modelos. En esta sección se van a estudiar métricas enfocadas a problemas de clasificación. Dentro de esto primero se va a definir una matriz de confusión.

### 2.2.4.1. Matriz de confusión

Esta es una herramienta que permite la visualización del desempeño de los algoritmos de clasificación, permitiendo observar la cantidad de registros que fueron clasificados correctamente en cada categoría y aquellos que se clasificaron de manera incorrecta. Para el caso de un problema de clasificación binaria esta matriz tiene 4 campos que se muestran en la Figura 2.4. Donde los valores colocados en

$a$  y  $d$  corresponden a aquellos que fueron clasificados con la etiqueta correcta y los otros campos  $b$  y  $c$  corresponden a los registros etiquetados de forma errónea.

		Predicted	
		Negative	Positive
Actual	Negative	$a$	$b$
	Positive	$c$	$d$

Figura 2.4: Matriz de confusión.

Esta herramienta se utiliza para estimar el desempeño real que va a tener el algoritmo en la operación real y así seleccionar el algoritmo que mejor se ajusta al problema. Para el caso específico del proyecto es útil para seleccionar el algoritmo y para realizar la evaluación técnica y económica del proyecto.

A partir de la matriz de confusión se desprenden otras métricas como las siguientes:

- **Accuracy:** Esta métrica corresponde al desempeño global del sistema y se define de la siguiente forma:

$$accuracy = \frac{c + d}{a + b + c + d} \quad (2.2)$$

Esta métrica es utilizada en el proyecto para determinar el porcentaje de casos que se clasifican de forma correcta, tanto aquellos que tienen problemas técnicos, como aquellos que no los tienen.

- **Precisión:** Esta métrica corresponde al porcentaje de registros que fueron clasificados correctamente de los que fueron etiquetados como positivos y se define de la siguiente forma:

$$precision = \frac{d}{b + d} \quad (2.3)$$

Esta métrica es relevante en el proyecto, ya que explica qué porcentaje de clientes que son clasificados con problemas técnicos, realmente tienen problemas técnicos.

- **Recall:** Esta métrica corresponde al porcentaje de registros positivos que fueron efectivamente capturados por el modelo. También es llamada tasa de verdaderos positivos.

$$recall = \frac{d}{c + d} \quad (2.4)$$

Esta métrica explica qué porcentaje de los clientes con problemas técnicos de la cartera de clientes, son realmente capturados por el modelo.

- **F1-Score:** Esta métrica corresponde a la media armónica de la precisión y el recall. Donde esta se calcula de la siguiente forma:

$$F1 = \frac{n}{\sum_i \frac{1}{x_i}} = \frac{2}{\frac{1}{P} + \frac{1}{R}} \quad (2.5)$$

Finalmente,

$$F1 = 2 * \frac{precision * Recall}{precision + recall} \quad (2.6)$$

Esta métrica es utilizada para seleccionar que modelo tiene mejor desempeño tanto en precisión como en recall. Ya que la media armónica penaliza de mayor forma que la media aritmética si alguno de los valores es muy bajo.

#### 2.2.4.2. Área bajo la curva ROC

El área bajo la curva roc (AUC-ROC) es otra métrica que se desprende de la matriz de confusión, pero en la cual se observan diferentes puntos de corte para la probabilidad del modelo. Para entender esta métrica se van a definir las siguientes métrica en base a la figura 2.4:

- **Tasa de Verdaderos positivos:** Esta métrica se define anteriormente bajo el nombre de recall y corresponde a las muestras positivas capturadas por el modelo.
- **Tasa de Falsos positivos:** Esta métrica corresponde a la tasa de registros negativos que fueron clasificados como positivos y se define de la siguiente forma:

$$FP = \frac{b}{a + b} \quad (2.7)$$

Entonces el área bajo la curva roc corresponde al área del gráfico con la tasa de falsos positivos en el eje x y la tasa de verdaderos positivos en el eje y.

Esta métrica es muy útil para definir que algoritmo tiene mejor desempeño cuando la data es desbalanceada.

#### 2.2.4.3. Prueba de Kolgomorov-Smirnov

La prueba de Kolgomorov-Smirnov o prueba KS es un test no paramétrico de bondad de ajuste y se utiliza para medir la distancia entre 2 distribuciones [Smirnov, 1939]. Esta es utilizada en el proyecto para la selección de variables que se introducen al modelo, a partir de la distancia que tienen las distribuciones de una variable para clientes con problemas técnicos versus aquellos que no tienen problemas técnicos. Entonces, se espera que aquellas varia-

bles que tienen una mayor distancia pueden distinguir de mejor manera el comportamiento.

# Capítulo 3

## Planteamiento estratégico y análisis de la situación actual

### 3.1. Posicionamiento estratégico

Para la definición del posicionamiento estratégico se recurre al modelo Delta [Hax and Wilde, 2003], el cual se puede observar en la Figura 3.1. El modelo Delta está compuesto por 3 ejes: Mejor producto, Solución integral con el cliente y Lock-in sistémico (Consolidación del sistema).

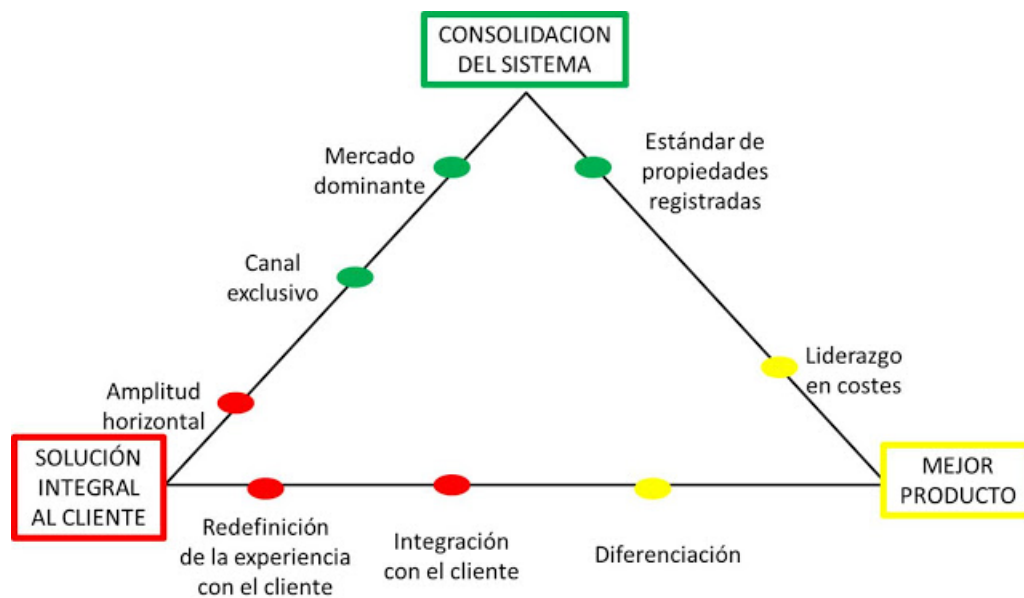


Figura 3.1: Modelo Delta de Hax.

De acuerdo al modelo anteriormente presentado, VTR tiene un posicionamiento estratégico de mejor producto por diferenciación, esto significa que busca ventajas competitivas a partir de la innovación en los procesos y a partir de esto entregar un

servicio único y de calidad al cliente, lo que se ve reflejado en la visión de la empresa que se mencionó anteriormente en la introducción. Asimismo, los factores relevantes en la calidad del servicio que son importantes para diferenciarse son: la continuidad del servicio, como ausencia de cortes y/o intermitencias; el cumplimiento de la velocidad contratada por el cliente; y una respuesta rápida y eficaz a las problemáticas que pueda tener el cliente. Un aspecto destacable para respaldar el posicionamiento estratégico, es el bajo costo de cambiarse de compañía, ya que al no brindar este servicio con la calidad que espera el cliente éste se puede cambiar a otra compañía que prometa y cumpla, a pesar de que hay mayor o menor competencia dependiendo del sector en que se encuentre el cliente.

### **3.2. Modelo de negocios**

El modelo de negocios permite llevar a la práctica el planteamiento estratégico de la empresa, como también permite representar de forma resumida la forma cómo la empresa genera valor a los clientes, cuáles son los clientes y cómo se logra generar ingresos.

Para esto se describe el modelo de negocios, según el marco de referencia de Kaggerman [Johnson et al., 2008], que se muestra en la Figura 3.2 y se describe posteriormente:

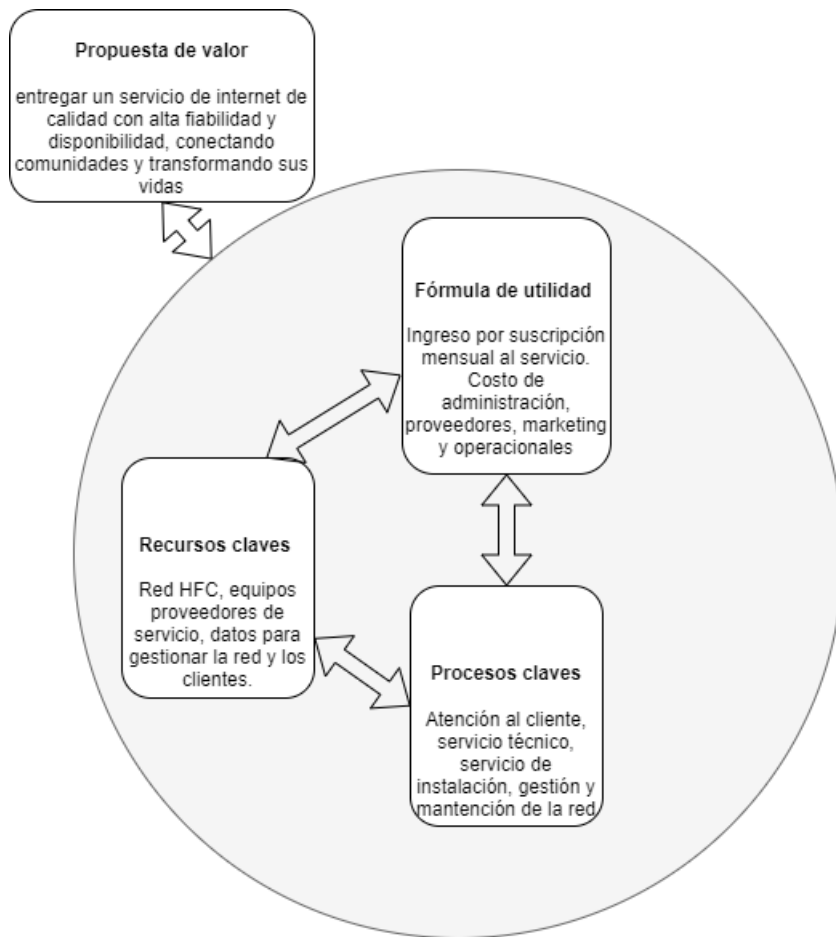


Figura 3.2: Modelo de negocios de VTR.

- **Propuesta de valor:** VTR tiene como clientes personas naturales y empresas pequeñas y medianas en menor volumen. La propuesta de valor hacia los clientes consiste en entregar un servicio de internet de calidad con alta fiabilidad y disponibilidad, conectando comunidades y transformando sus vidas.
- **Fórmula de Utilidad:** La fórmula de utilidad de VTR en la línea de negocios de internet fija es la misma que la de cualquier compañía de telecomunicaciones. Se obtiene ingresos por una suscripción mensual a los clientes que contratan el servicio y los costos son principalmente administración, conexión del servicio a los proveedores comerciales y operacionales de mantenimiento de la red, entre otros.
- **Recursos claves:** Los recursos claves que permiten generar ventajas competitivas son la red HFC que corresponde al cableado de fibra óptica y cable coaxial, también los equipos que proveen el servicio y finalmente los datos juegan un papel clave como recurso, ya que permiten una mejor gestión de las redes y de los clientes.

- **Procesos claves:** Los procesos claves que permiten generar valor son los procesos de cara al cliente, como atención al cliente, servicios técnicos y servicios de instalación. También, los procesos operacionales de gestión y mantenimiento de la red que permite transportar el servicio entregado.

### 3.3. Diagnóstico de la Situación Actual

Como se menciona en el primer capítulo, a mediados de Marzo se declara cuarentena obligatoria para varias regiones de Chile, lo que provoca un aumento masivo en el tráfico de internet, ello trae consigo no solo un mayor uso de internet por parte de los clientes, sino también, una mayor exigencia por parte de estos, como consecuencia, aumentan la cantidad de reclamos en el servicio de internet.

El cliente ahora utiliza internet en diferentes horarios y con una exigencia mayor, ya que hace la vida cotidiana a partir de esta herramienta, lo cual golpea a las empresas de telecomunicaciones que no estaban preparadas para algo así y especialmente golpea a las empresas con mayor cantidad de clientes activos como es VTR.

Las métricas que utiliza VTR para medir la experiencia del cliente son dos: la primera son las llamadas ingresadas a la plataforma técnica, es decir, las llamadas que al ingresar seleccionan la opción de problema técnico (aun cuando la llamada puede o no ser contestada), específicamente a la plataforma técnica de internet; y la segunda, los truck rolls que corresponden a los servicios técnicos que van al domicilio. Como se puede observar en la Figura 3.3 la tasa de llamados por motivos técnicos ha aumentado más de un 100% entre los meses previos a la cuarentena y los meses posteriores, tasa que se ha mantenido alta a pesar de los trabajos de aumento de capacidades que finalizaron se forma temprana, en estos también se observa que han decrecido en el último periodo, pero debido al cansancio de algunos clientes con llamar a VTR. La alta tasa de llamados ha generado problemas en la relación con el cliente, ya que durante el periodo peak tuvo al call center colapsado gran parte del tiempo, porque los clientes no eran atendidos por VTR o los tiempos de espera eran mayores de lo habitual, a pesar de que las llamadas han ido a la baja en el último periodo, teniendo los últimos meses de septiembre y octubre en valores normales, la tasa de desconexiones voluntarias del servicio de internet fijo ha ido al alza. Se puede observar en la Figura 3.4 que, a pesar de que en los primeros meses de la pandemia la tasa no era tan alta, esto se debe a que no es tan fácil para los clientes pedir una desconexión del servicio en pleno peak de la pandemia. Sin embargo, a partir del mes de junio se ve un incremento en esta tasa el que se ha sostenido durante los siguientes meses y teniendo su peak en el mes de octubre, siendo casi el doble



de desconexiones que en periodos normales.

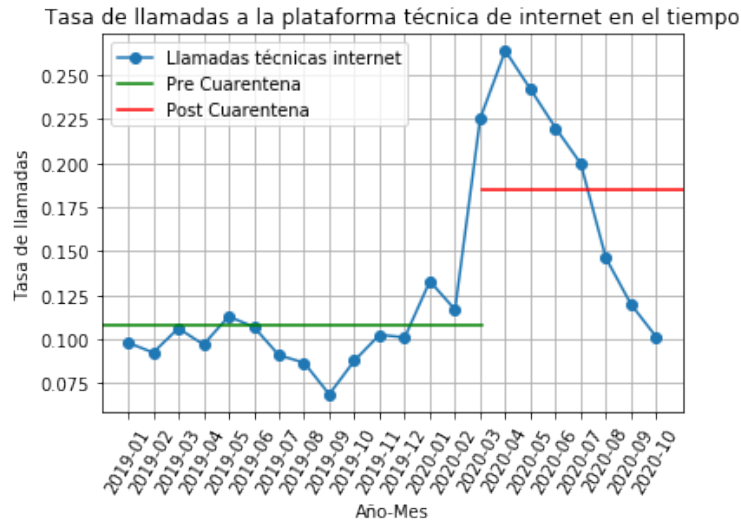


Figura 3.3: Evolución de la tasa de llamadas por motivos técnicos.

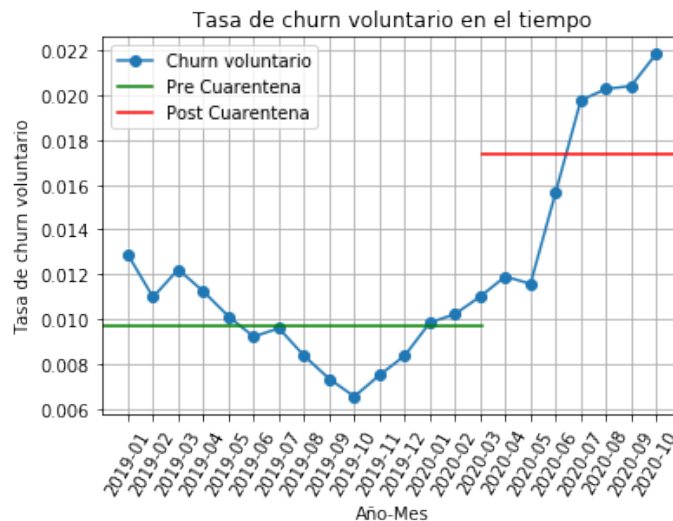


Figura 3.4: Evolución de la tasa de fuga del servicio de internet.

### 3.3.1. Problema Identificado / Oportunidad identificada

Para el problema identificado anteriormente, que corresponde a la alta tasa de desconexiones del servicio de internet, se busca saber la causa raíz y a partir de esto buscar una oportunidad de mejorar la situación. La metodología utilizada para buscar la raíz del problema es un árbol why que se muestra en la 3.5.

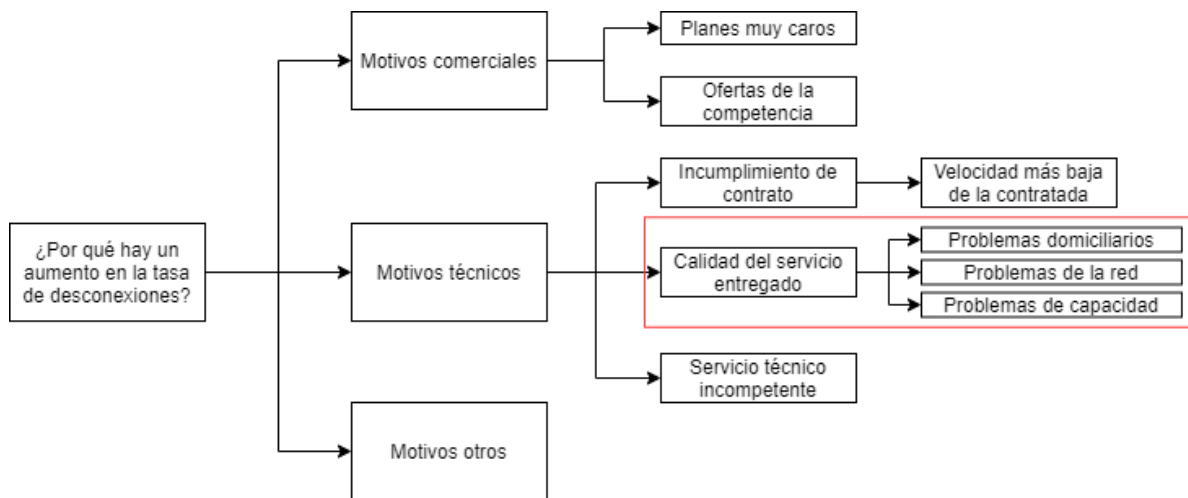


Figura 3.5: Árbol Why - Aumento desconexiones.

Para entender cuál es la razón del aumento de desconexiones se puede observar un árbol why en la Figura 3.5, donde la primera capa se desprende los motivos de las desconexiones de internet por parte de los clientes, entre estos están los motivos comerciales, motivos técnicos y otros. Desde VTR se consulta a un subconjunto de clientes que se fugan, al consultarles los motivos de su desconexión, el 60 % dice que es por motivos técnicos y malestar con la calidad del servicio entregado, la segunda gran parte se debe a motivos comerciales y la menor parte se debe a otros motivos. Al abrir cuales son los motivos técnicos que más influyen son el incumplimiento del contrato (refiriéndose a que la velocidad real es más baja que la velocidad contratada), el servicio técnico incompetente, es decir, que van técnicos a reparar los problemas del hogar y estos no se reparan o la falta de respuesta desde el call center, y, finalmente, la calidad del servicio de internet vista como la intermitencia del servicio el cual no es constante, donde hay caídas del servicio, días que no funciona, etc. Esto se puede deber a problemas de la planta externa, a problemas del cableado domiciliario o problemas de capacidad de la red y también problemas puntuales que tienen una ocurrencia baja, tales como cortes de fibra, cortes de luz, interferencias puntuales, etc.

Del árbol Why, se desprende que el aumento de desconexiones de internet se explica en gran parte por problemas con la calidad del servicio entregado, lo que es respaldado por bibliografía que muestra la influencia de la calidad del servicio en la decisión de cambiarse de compañía proveedora de internet fijo [Dwivedi et al., 2010]. También por la correlación que existe entre el aumento de desconexiones y el aumento de llamados por motivos técnicos, siendo este último, otro síntoma de la mala experiencia del cliente con el servicio.

Entonces, el problema raíz que está detrás del problema de negocio de la fuga de clientes corresponde a que la calidad del servicio entregada al cliente no es suficiente o no se ajusta a la expectativa del cliente y con esto la experiencia que

tiene el cliente se ve degradada y decide cambiar de proveedor de internet, para buscar alternativas sobre cómo solucionar este problema se utiliza la metodología del árbol How que se muestra en la Figura 3.6.

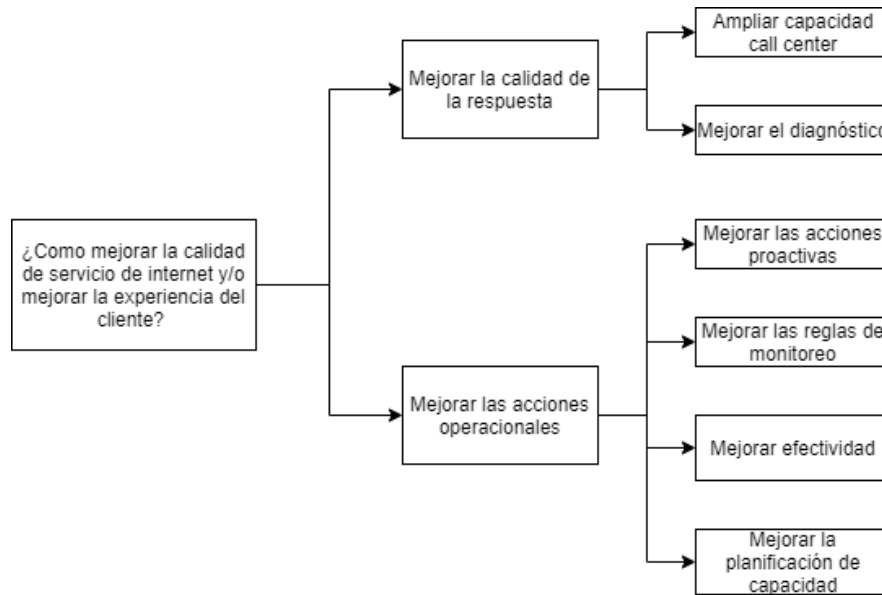


Figura 3.6: Árbol How - Mejorar calidad de servicio.

De la Figura 3.6 se desprenden 2 dimensiones de la calidad de servicio, la primera es la calidad de respuesta de cara al cliente, la que se abre en ampliar la atención al cliente (ampliando el call center) y mejorar el diagnóstico de los problemas del cliente, mejorando así la tasa de resolución de problemas. Por otro lado, es mejorar las acciones operacionales, donde se abren varios frentes, ya que las acciones operacionales pueden ser en la red, capacidades o en el domicilio. Para el caso específico de este proyecto se considera la solución referente a las acciones operacionales. Las opciones de este nodo se desprenden de la siguiente manera:

- **Mejorar planificación de capacidades:** Esta opción tiene relación a anticiparse a los peak de capacidades que pueden haber y no esperar que la red colapse para aumentar la capacidad.
- **Mejorar efectividad:** Con el fin de garantizar que la red tenga la calidad necesaria para entregar el servicio sin problema, es necesario garantizar que las acciones que se realizan tengan una alta efectividad, tanto que solucionen el problema que deben solucionar como que no generen un impacto negativo en otra área, para esto es necesario monitorear estas acciones y ver la forma de mejorarlas en caso de que no funcionen como deberían.
- **Mejorar acciones proactivas:** Esto tiene relación a anticiparse al reclamo del cliente, tiene relación a entender cuando un cliente tiene problemas y poder repararlo de forma temprana.

- **Mejorar las reglas de monitoreo:** Hoy día la forma de monitorear la experiencia del cliente es a partir de las llamadas que realiza a la plataforma técnica y a las métricas de desempeño que entrega su cable módem y a partir de estas métricas se observan algunas y en base a umbrales definidos por el proveedor se determina si un cliente está bien o está mal. El principal problema con el monitoreo sucede cuando las reglas no están bien definidas, ya sea porque son muy estrictas o muy livianas; por un lado cuando son muy estrictas, el cliente puede estar sufriendo y esto no es visible para la operación del servicio al cliente; por otro lado cuando las reglas son muy livianas, el número de clientes que se ven en mal estado es muy grande y difícilmente se puede apuntar con acciones directas.

En base a las hipótesis mencionadas anteriormente, se entienden estas como complementarias para mejorar la calidad del servicio. Sin embargo, el proyecto tiene foco en la hipótesis de mejorar las acciones proactivas, debido a la ausencia de acciones de este tipo en la operación habitual.

La idea de realizar acciones proactivas es evitar el malestar del cliente, evitar que se caiga el servicio o que tenga problemas de experiencia de algún tipo, manteniendo el servicio en niveles óptimos y evitando algún reclamo por parte del cliente. La relevancia de evitar el reclamo es mantener la experiencia del cliente impecable y de esta manera evitar que decida cambiarse de compañía proveedora de internet fijo como se ha observado hasta el momento, cuando la mala experiencia de los clientes y su inconformidad con el servicio, ha provocado que los clientes tomen esta decisión. Lo favorable de realizar acción proactiva al domicilio es que esto no es invisible a los ojos del cliente, ya que es notificado que a la compañía le es relevante que su servicio se mantenga en buenas condiciones y como adicional, se certifica que todos los servicios contratados por el cliente tales como telefonía fija y/o televisión digital estén funcionando de forma correcta. Siendo esta acción tanto para mantener la calidad del servicio, como para fidelizar a los clientes que sean tocados.

### **3.3.2. Arquitectura de Procesos AS IS**

Para entender qué procesos hay que intervenir o crear en la empresa, se recurre a la herramienta de los macroprocesos [Barros, 2015]. VTR tiene una cadena de valor unificada, donde los diferentes servicios comparten gran parte de la cadena de valor y difieren en algunos procesos únicamente.

El proceso a intervenir está situado en la macro 1 que corresponde a la cadena de valor de VTR, dentro de esta macro se profundiza en la administración y relación con los clientes, debido a que en este macroproceso se tiene 3 grandes procesos:

el estudio de mercado y la planificación de ventas, venta y atención a clientes de internet y decisión de satisfacción de requerimientos. En el marco de esto, es necesario adentrarse en la venta y atención a clientes de internet, donde se encuentra la canalización de venta y consulta de clientes de internet, preventa y cotización del servicio de internet fijo y servicio de atención de clientes de internet. Este último, corresponde a los procesos que hay que intervenir, ya que corresponden a los procesos que se encargan de monitorear los niveles del servicio al cliente, y decidir qué acciones se deben tomar para mejorar tanto comercial como técnicamente; aunque el foco del proyecto es la parte técnica y es donde actualmente se tiene un proceso de resolución de incidentes de los clientes, que se ejecuta de forma reactiva, cuando el cliente reclama por el problema que experimenta.

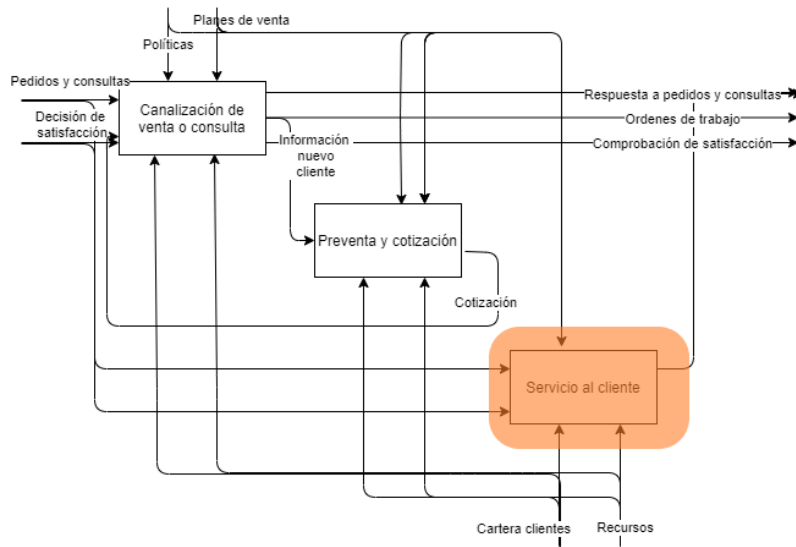


Figura 3.7: Venta y atención de clientes de internet fijo.

### 3.3.3. Modelamiento Detallado de Procesos AS IS

En la Figura 3.8 se muestra el proceso de resolución de incidentes técnicos del cliente.

Este proceso comienza cuando el cliente llama al call center de VTR. Entonces, el cliente elige qué plataforma desea escoger, donde tiene las opciones de comercial para resolver problemas de facturación u otros relacionados, la plataforma de venta para contratar servicios, la plataforma de desconexión para darse de baja de algún servicio o la plataforma técnica para indicar que tiene algún problema en el servicio. Al iniciar esto el ejecutivo del call center sigue un protocolo para ayudar a solucionar el problema al cliente, finalmente, si no se reestablece el servicio se genera un ticket para un Truck roll (servicio técnico domiciliario).

Por otro lado, este servicio al cliente es únicamente reactivo, es decir, se acciona a partir de que el cliente tenga un malestar con el servicio y decida llamar a VTR para solucionarlo, actualmente no existe mecanismo alguno para realizar esta acción de forma proactiva, mejorando la experiencia del cliente y asegurando una calidad del servicio.

Esto provoca que el cliente tenga que sentir malestar con el servicio y sufrir alguna caída para poder solucionar su problema, como también provoca que VTR no pueda balancear la carga tanto de sus canales de atención, como de los técnicos que realizan visitas domiciliarias estando desprovistos para momentos peak que ocurran, como fue al inicio de la pandemia.

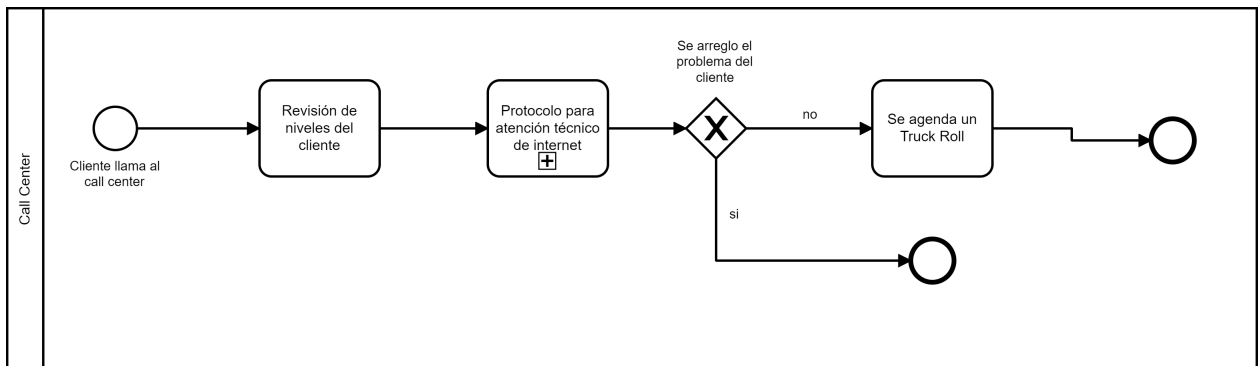


Figura 3.8: Proceso de resolución de incidentes técnicos.

### 3.4. Cuantificación del Problema

Los números que justifican este problema son citados al inicio del capítulo, para recordar comenzaremos observando la Figura 3.4, donde se muestra que a partir del mes de junio las desconexiones voluntarias aumentan de un 1.2% a un 1.6%, es decir un 30% más de desconexiones. La cartera total aproximada de clientes de internet es de 1.3 Millones de clientes, de este total, el 0.4% extra de desconexiones corresponde a 5.200 clientes que se fugan. Al observar los meses posteriores a junio se puede ver que el porcentaje de fuga de clientes no decrece, sino que aumenta llegando al 2% de clientes que se fugan estabilizándose cerca de este número y manteniendo este alto nivel de fuga durante los siguientes meses. Tener una tasa de fuga voluntaria constante cercana al 2% de clientes, significa un 66% más de fuga de lo que era un mes habitual antes de la pandemia y traducido en número de clientes es cercano a los 10.000 clientes. Si se toma el plan más barato de internet fijo en VTR, que tiene un valor cercano a los 20.000\$ pesos chilenos. Se obtiene un total de 200 Millones de pesos que VTR deja de percibir en un mes por la fuga extra de clientes.

Tabla 3.1: Cuantificación del problema

	<b>MES</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Total</b>
<b>Tasa Cte.</b>	Extra clientes fugados	10.000	10.000	10.000	10.000	10.000	50.000
	Acumulado clientes	10.000	20.000	30.000	40.000	50.000	50.000
	Dinero no percibido (pesos chilenos)	200 M	400 M	600 M	800 M	1.000 M	<b>3.000 M</b>
<b>Tasa Crec.</b>	Extra clientes fugados	10.000	11.000	12.000	13.000	14.000	60.000
	Acumulado clientes	10.000	22.000	34.000	46.000	60.000	60.000
	Dinero no percibido (pesos chilenos)	200 M	440 M	748 M	1.012 M	1.320 M	<b>3.720 M</b>
<b>Tasa Decrec.</b>	Extra clientes fugados	10.000	9.000	8.000	7.000	6.000	40.000
	Acumulado clientes	10.000	19.000	27.000	34.000	40.000	40.000
	Dinero no percibido (pesos chilenos)	200 M	380 M	540 M	680 M	800 M	<b>2.600 M</b>

En la Tabla 3.1 se observa la cantidad de clientes extra que se fugan por mes y el acumulado a 5 meses, para 3 diferentes escenarios: en el primero, la cantidad de clientes extra fugados es constante, en el segundo, aumenta de forma constante una cantidad de 1.000 clientes mensuales y el tercero decrece de forma constante en 1.000 clientes mensuales. Según estos supuestos, el costo en pesos chilenos está entre los 2.600 y 3.720 Millones de pesos chilenos que se dejan de percibir. Esta cifra muestra la importancia de realizar acciones remediales que impidan su ocurrencia en el futuro.

# Capítulo 4

## Propuesta de rediseño de procesos

Como se establece en el capítulo anterior, se va intervenir el proceso de resolución de incidentes técnicos del cliente a partir del diseño de un proceso de atención proactiva a clientes con problemas en el servicio de internet. En este capítulo se explicará la solución detallada que se propone con las direcciones de cambio, el alcance que tiene la solución, los métodos que se utilizan para solucionar el problema y la integración con los procesos de la empresa.

### 4.1. Direcciones de cambio y alcance

En esta sección se definen las direcciones de cambio, las cuales tienen como objetivo guiar y establecer qué procesos se tienen que implementar. Para entender la direcciones de cambio se van a estudiar las variables de diseño [Barros [2010]]. A continuación, se muestra la revisión de las direcciones de cambio que impacta el proyecto.

La Tabla 4.1 muestra las variables de anticipación y de coordinación. Actualmente, las acciones de aseguramiento de la calidad del servicio en el domicilio del cliente se hacen de forma reactiva, es decir, se espera que el cliente reclame por el servicio y contacte a VTR para iniciar un ticket de servicio domiciliario. Variable de anticipación: el proyecto propone que se realice de forma proactiva, identificando a los clientes con problemas en el servicio, priorizándolos y contactando a aquellos más críticos para proponer una visita proactiva.

La variable de coordinación considera 2 variables en las que influye el proyecto, estas son reglas y jerarquía.

Como se observa en la Tabla 4.1, las reglas para entender si un cliente tiene problemas con el servicio son a partir de un conjunto acotado de métricas y éstas se encuentran en umbrales definidos por el proveedor en forma individual; sin tomar en cuenta la interacción entre ellas, considerando tan solo un subconjunto de los



Tabla 4.1: Variables de diseño

Variable de diseño	Actual	Propuesto
<b>a. Anticipación</b>		
Modelo predictivo de fallas en el servicio	No	Modelo predictivo
<b>b. Coordinación</b>		
Reglas	Establecidas por el proveedor	Establecidas de forma estadística
Jerarquía	Siempre	Ocasionalmente

problemas que puede tener el cliente y dejando algunos de lado. El proyecto propone que las reglas estén establecidas de forma estadística, abarcando la mayor cantidad de los problemas que sufren los clientes.

La otra variable que muestra la Tabla 4.1 es la de jerarquía, ella indica que las acciones se tienen que tomar en forma reactiva, son hoy las únicas que se realizan de cara al cliente; en cambio, con el proyecto se propone que además de las acciones reactivas se realicen acciones proactivas, permitiendo así mantener los niveles de servicios correctos.

Éstas corresponden a las 2 variables principales en que se propone cambio con el proyecto. Actualmente, que las acciones sean activadas en forma reactiva sobre la base del reclamo del cliente, degrada la experiencia que el cliente tenga con la empresa; por el contrario, el proyecto busca llegar a clientes que teniendo problemas, aún no hayan reclamado.

## 4.2. Propuesta de solución

La solución que se propone interviene en el proceso de resolución de incidentes técnicos de internet fijo, a partir del diseño y creación de un proceso de atención proactiva a los clientes que tengan problemas con el servicio de internet y no hayan reclamado, ya sea porque no se han dado cuenta o la falla es incipiente y no afecta al cliente de tal forma como para que haya decidido reclamar.

Para el funcionamiento del proyecto se deberá intervenir algunas áreas. La primera es el área de BI que se encarga de los desarrollos necesarios para que el modelo de priorización de clientes de internet sea ejecutado y cargado en las bases de atención al cliente. La segunda área es la de clientes, la que se encarga del contacto con el cliente, la que en conjunto con el área técnica tienen que decidir la cantidad de clientes a ser contactados por semana. Luego de decidir qué clientes contactar y realizado este contacto proactivo indicando que se enviará a un técnico para asegurar que el servicio esté en los niveles óptimos y no se tenga problemas

con la experiencia del servicio. A los clientes que acepten, se les agenda la visita; entra al proceso el área técnica, enviando al técnico domiciliario, quien se encarga de arreglar el incidente técnico y de certificar el correcto funcionamiento de todos los servicios que tiene el cliente.

Lo anterior, con el propósito de mantener una buena experiencia de servicio por parte del cliente, evitando que posteriormente tenga algún problema y deba contactar a VTR para darle aviso. También permite balancear la carga de los canales de atención y de los técnicos domiciliarios.

La priorización de clientes se realizará aplicando un modelo para determinar qué clientes tienen problemas con el servicio de internet, a partir de las métricas de desempeño que se recolectan desde los cable módem y permiten priorizar a los clientes con problemas críticos, considerando las llamadas históricas de los clientes a la plataforma técnica de internet.

#### **4.2.1. Modelamiento Detallado de Procesos TO BE**

En la Figura 4.1 se muestra el proceso que se quiere implementar para realizar las visitas domiciliarias proactivas. Este proceso comienza con las consultas de la plataforma llamada Servassure a los cables módem de los clientes recolectando las métricas de desempeño de estos. Posteriormente, el área de BI ejecuta el modelo de detección de clientes con problemas de internet y se guardan las etiquetas en la base de datos. Luego, se priorizan aquellos clientes que tienen mayor valor, es decir, que tienen más productos contratados, y se cargan esos RUT al equipo de clientes. Finalmente, el equipo de clientes planifica la cantidad de clientes a ser contactados en la semana y realiza el envío de mensajes a los clientes con problemas críticos de forma proactiva. En este mensaje se informa que los niveles del cliente no estarían en el óptimo para el nivel de servicio de VTR y se propone una visita técnica al domicilio para dejar los servicios en su correcto funcionamiento. En este punto, el cliente toma la decisión de aceptar o rechazar la visita técnica; en caso de rechazarla, se carga a la plataforma de clientes por si se produce un posterior contacto por parte del cliente. En caso que el cliente acepte la visita técnica, ésta se agenda y el equipo técnico realiza la visita certificando todos los servicios.

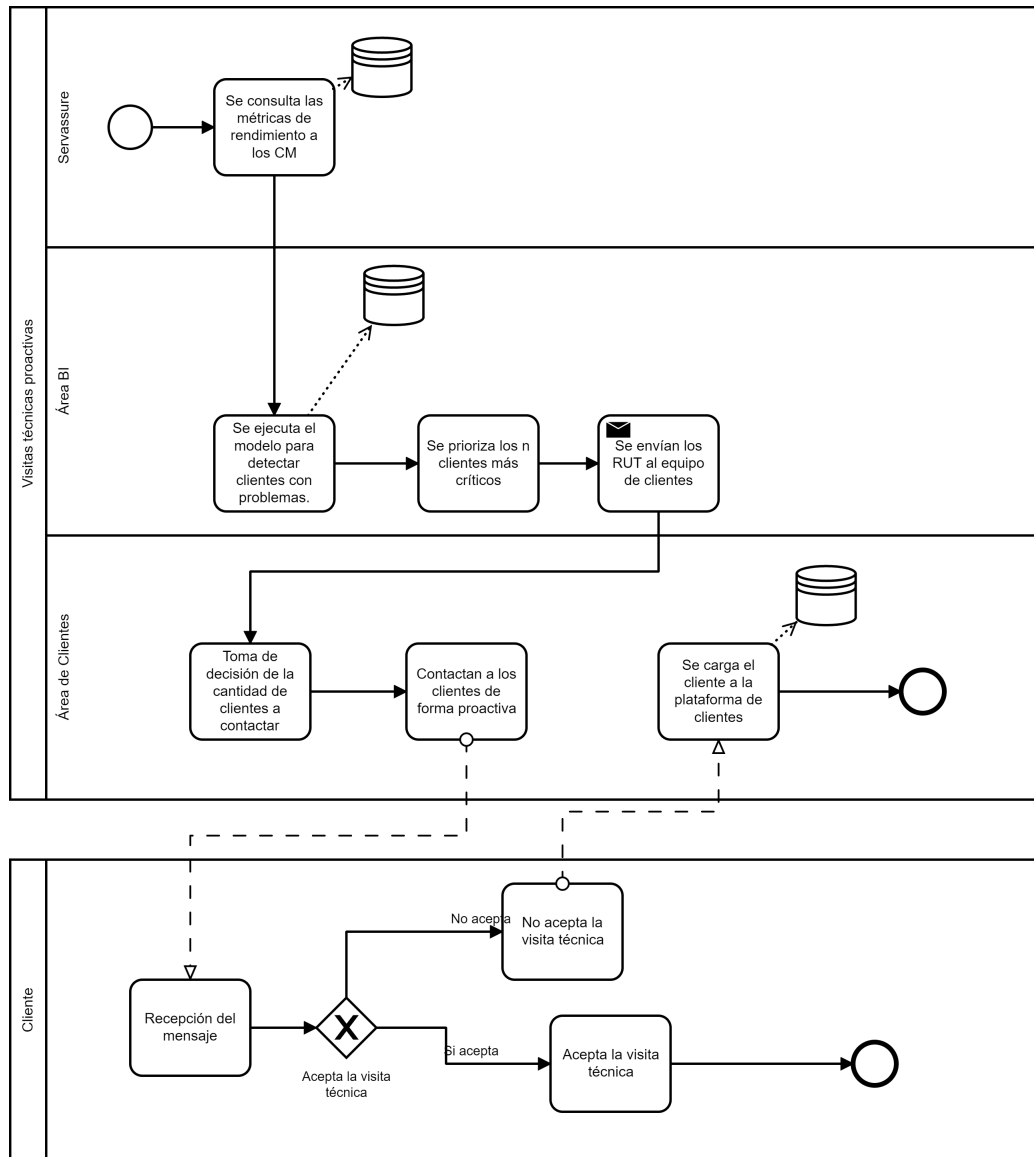


Figura 4.1: Propuesta proceso de visitas domiciliarias proactivas.

#### 4.2.2. Diseño de Lógica de Negocios

Como se puntualiza en capítulos anteriores el proyecto consiste en detectar cuándo un cliente tiene problemas con el servicio de internet, aunque no reclame o no advierta su problema. Para realizar dicha tarea hay que definir la aproximación y los supuestos que se realizan para detectar un problema en el servicio.

Para realizar las tareas antes mencionadas se tiene 3 dataset a disposición: el primero tiene relación con las llamadas de los clientes por motivos técnicos; el segundo, son las métricas de desempeño de la red que se extraen de los cable modem; por último, es el stock de los clientes activos. A continuación, en las Tablas 4.3, 4.2, 4.4 se resumen las variables que se tienen disponibles para realizar el

modelamiento.

La llamada del cliente a la plataforma técnica se tiene como aproximación, es decir, el cliente llama a la plataforma técnica de internet porque tiene un problema con el servicio de internet, siente que algo no está funcionando como debería o no tiene servicio en el momento que llama. La complicación de esta variable es el ruido intrínseco que posee, porque hay clientes que llaman y que no necesariamente tienen problemas en el servicio o el problema no se debe a las métricas del cable módem, sino a dificultades de utilización, masivos, entre otros. El segundo caso es de clientes que sí tienen problema con el servicio, pero no llaman debido a que no usan el servicio en los horarios en que éste se está presentando, o no detectan el deterioro que puede estar teniendo el servicio. Frente a esta situación, se entiende que aquellos clientes que tienen una alta tasa de llamados y que tienen unas métricas parecidas son clientes con problemas en el servicio; por el contrario, clientes que tienen una tasa baja de llamados y métricas parecidas son clientes que no tienen problemas con el servicio.

Un ejemplo para entender la lógica del supuesto anterior y corroborar su veracidad. Al observar la Figura 4.2, que corresponde a la variable de señal sobre ruido en subida (SNR UP) vs la tasa de llamadas, se observa que al tener una señal sobre ruido más baja la tasa de llamados aumenta significativamente. Por ejemplo, el grupo que tiene esta métrica por debajo de los 20 dB tiene una tasa de llamados de 0.78 % que es casi 3 veces mayor a aquellos que tienen esta métrica en buenos valores. Tener un SNR UP menor a 20 entrega una mala calidad de internet. Sin embargo, ni el 1 % de los clientes en esta condición, llama a la plataforma técnica en el día que se presenta. Por otro lado, se observa que quienes tienen esta métrica con valores mayores a 37 dB, igual llaman a la plataforma técnica, aunque en este caso podría darse que tuvieren problemas con otra variable; si bien esta vista es univariada, sirve como ejemplo para entender la naturaleza del problema. Las razones de que no llamen pueden ser: que el problema se está presentando en horarios en que el cliente no utiliza el servicio, o el uso del cliente no permite detectar el problema y/o el problema es incipiente y el cliente no lo detecta aún, siendo probable que lo haga después.

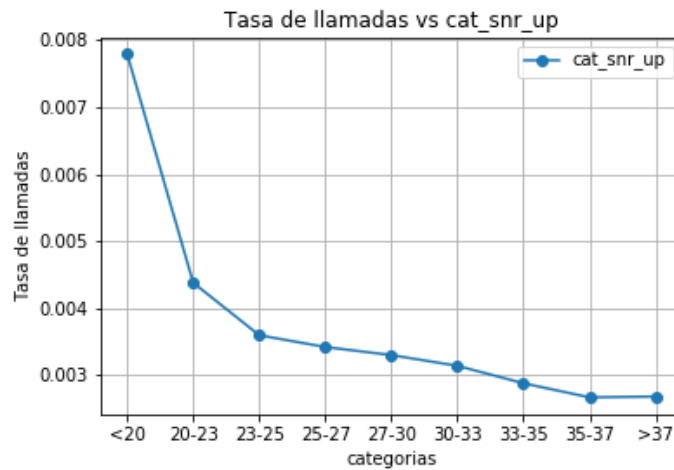


Figura 4.2: Ejemplo del supuesto de llamadas técnicas en la detección de clientes

Por las razones expuestas, es posible que el error de clasificación sea alto lo cual no significa que el modelo de detección de problemas esté errado, porque se utiliza una aproximación al problema. Lo que importa en la medición del modelo, es que se pueda encontrar umbrales bajo los cuales el servicio se vea degradado.

### 4.3. Diseño del modelo de detección de problemas del servicio de internet fijo

La primera etapa de la metodología CRISP-DM es el entendimiento del negocio donde se definen los objetivos y requerimientos desde la perspectiva del negocio. Además se analiza la situación actual y el problema de negocio que se desea resolver. Esta etapa no se abarcará en esta sección, ya que la descripción de la situación actual y el análisis de problema se describe en las secciones anteriores.

A modo de resumen, se quiere determinar qué clientes tienen problemas con el servicio de internet fijo a partir de las llamadas a la plataforma técnica de VTR y de las métricas de desempeño de los cables módem. Entonces, el supuesto del modelamiento es el siguiente: si clientes que tienen métricas de desempeño parecidas a un grupo de clientes con una tasa de llamados mucho más alta que la cartera, implica que estos clientes tienen problemas con el servicio de internet, ya sea que pertenezcan o no, al grupo que llama. Por el contrario, si clientes tienen métricas de desempeño parecidas y una tasa de llamados mucho más baja que la cartera, se entiende que los clientes no tienen problemas de calidad del servicio; aunque pueden tener otros problemas, como de capacidades, por ejemplo, los que no forman parte del alcance del proyecto.

### 4.3.1. Entendimiento de los datos

La primera tarea de esta etapa es disponibilizar los datos con los que se va a trabajar, lo que se puede observar en las Tablas 4.2, 4.3, 4.4.

Tabla 4.2: Métricas de desempeño de los cable módem

<b>Variables</b>	<b>Descripción</b>
<b>MAC</b>	MAC del cable modem del cliente
<b>Fecha</b>	Fecha de medición
<b>Hora</b>	Hora de medición
<b>CER</b>	Porcentaje de paquetes con errores no corregidos
<b>CCER</b>	Porcentaje de paquetes corregidos
<b>SNR</b>	Señal sobre el ruido
<b>RXpower</b>	Poder de recepción
<b>TXPower</b>	Poder de transmisión
<b>T3</b>	Error T3
<b>T4</b>	Error T4
<b>Estado</b>	Estado del cable modem (offline, online, etc)

Tabla 4.3: Llamadas

<b>Variables</b>	<b>Descripción</b>
<b>Fecha</b>	Fecha en la que fue efectuada la llamada
<b>Rut Cliente</b>	Rut del cliente que efectua la llamada
<b>Plataforma entrada</b>	Plataforma por la cual ingresa
<b>Descripción</b>	Descripción del motivo de la llamada
<b>Duración llamada</b>	Duración de la llamada

Tabla 4.4: Cartera de clientes

<b>Variables</b>	<b>Descripción</b>
<b>Rut Cliente</b>	Rut del cliente
<b>MAC</b>	MAC del cable modem del cliente
<b>Velocidad</b>	Velocidad contratada
<b>Antigüedad</b>	Antigüedad del cliente
<b>Cuadrante</b>	Cuadrante que pertenece el cliente
<b>CMTS</b>	CMTS que pertenece el cliente

En la Tabla 4.2 se encuentran las métricas de desempeño recolectadas de los cable módem de los clientes, esta tabla tiene las MAC, la fecha, la hora de medición y las métricas de desempeño, tanto en upstream, como en downstream, es decir, la medición en entrada y salida del cable módem. La volumetría de esta tabla es de 15 columnas y 1.3 Millones de filas por hora. La Tabla 4.3 tiene la información de

las llamadas de los clientes por RUT, dando una aproximación de 5.000 llamadas diarias. Finalmente, la tabla que permite unir la información de los niveles con la información de las llamadas es la Tabla 4.4 que tiene la cartera completa de clientes con el RUT, la MAC de su cable módem, la velocidad contratada, entre otra información.

En la elección de la variable objetivo se utilizan únicamente las llamadas que ingresan a la plataforma técnica de internet con el objetivo de disminuir el ruido que tienen las llamadas de parte de los clientes. En la Figura 4.3 se observa la tasa de llamadas normalizadas respecto a diferentes rangos de la variable de señal sobre ruido en subida (SNR UP), donde se observa que la variable de llamadas técnicas de internet genera una mayor diferencia porcentual al variar este rango.

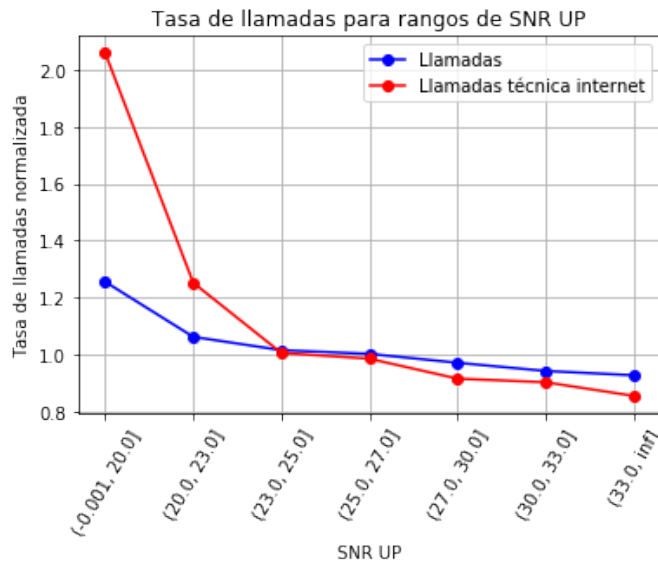


Figura 4.3: Comparación Llamadas vs llamadas técnicas de internet.

Como definición la variable objetivo del modelo son las llamadas que ingresan a la plataforma técnica de internet las cuales tienen una tasa diaria de 0.4 %, lo que induce a un problema de clases desbalanceadas. Debido a la volumetría de la Tabla 4.2 que es de 1.3 Millones de registros por hora (218 Millones de registros a la semana) y lo desbalanceada de las clases, se utiliza una granularidad diaria para los datos, en la que para cada variable disponible en la Tabla 4.2 se calcula el mínimo, máximo, promedio, varianza; se contabiliza la cantidad de horas que la variable toma valores fuera de rangos operacionales y fuera de rangos propuestos, como se observa en la Tabla 4.5 y también se contabiliza la cantidad de veces que la variable toma un valor nulo. Considerando estas agrupaciones, la cantidad de métricas disponibles (12) y el conteo de horas que el cable módem tiene en estado offline, online y otro estado, se obtienen 82 variables explicativas.

Para entender cada variable de forma univariada y su comportamiento con la va-

riable objetivo se generan agrupaciones de los datos por categorías indicadas por la operación y tablas descriptivas para ver la relación respecto a la tasa de llamados. Al realizar esta visualización, se observa el porcentaje de clientes que su cable módem entrega valores nulos, este porcentaje ronda el 0.5 % de clientes con valores nulos en algunas variables y el 2.5 % en otras. Por ejemplo, en la Figura 4.4 se observa que a mayor cantidad de paquetes con errores de subida se tiene una mayor tasa de llamados.

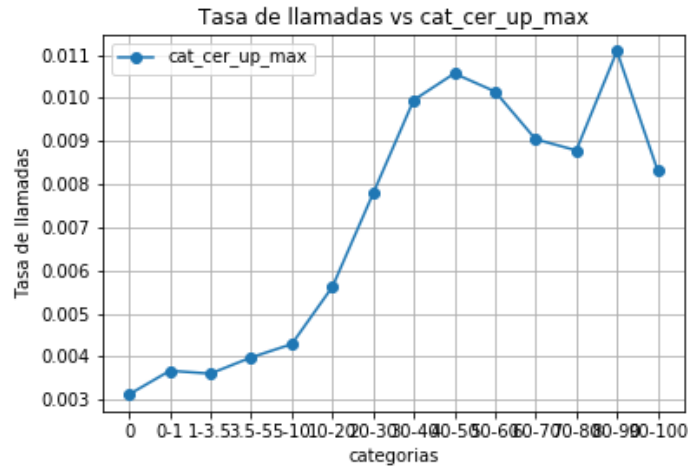


Figura 4.4: CER UP Max vs Llamadas técnicas.

Por otro lado, en la Figura 4.5 se observa que los clientes que tienen la variable T3 con un valor máximo mayor o igual a 2 durante el día, tienen un 66 % más de llamadas que aquellos clientes que tienen un valor 0.

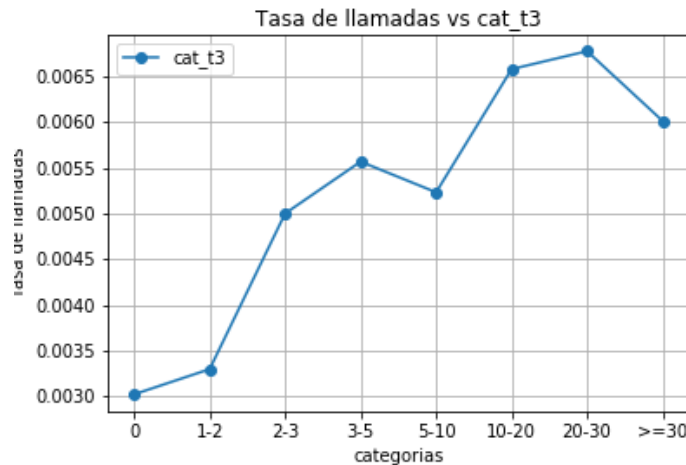


Figura 4.5: T3 Timeouts vs Llamadas técnicas.

Este tipo de gráficos nos indica cómo se va moviendo la tasa de llamados a medida que la variable toma diferentes valores y nos da el insight sobre los datos perdidos,



entregando una información relevante respecto a la degradación del servicio del cliente e indica que no están perdidos por mera casualidad. Por ello el tratamiento que se les da es diferente, no se puede imputar la media únicamente, sino debe explicitarse en el modelo que hay datos nulos para cada uno de los niveles. Ejemplo, si queremos explicitar los nulos que tiene la variable TXPOWER UP, la variable que indica los valores nulos en este nivel es NULL\_TXPOWER\_UP.

Posteriormente, para entender los datos que no son nulos, estos son filtrados y se realiza el mismo ejercicio anterior, con el fin de determinar umbrales individuales para las variables y contar la cantidad de horas fuera de los umbrales operacionales, comparándolos con los umbrales definidos por el proveedor y tomando una decisión sobre qué umbrales se van a contabilizar como fuera de rango. A modo de ejemplo, se muestra en la Figura 4.6 las categorías para la variable Rx power en upstream, es decir, el poder de recepción en subida para los cable módem, en el que se observa que el gráfico tiene forma de U, el proveedor y la operación indica que los valores normales de la variable están entre 14.5 y 15.5, pero al ver el gráfico se observa que la tasa de llamadas entre 15.2 y 15.5 es igual de alta que para valores mayores a 15.5, por lo que cada hora que la variable tome el valor fuera del rango [14.5, 15.2] se va a contabilizar en una variable del modelo.

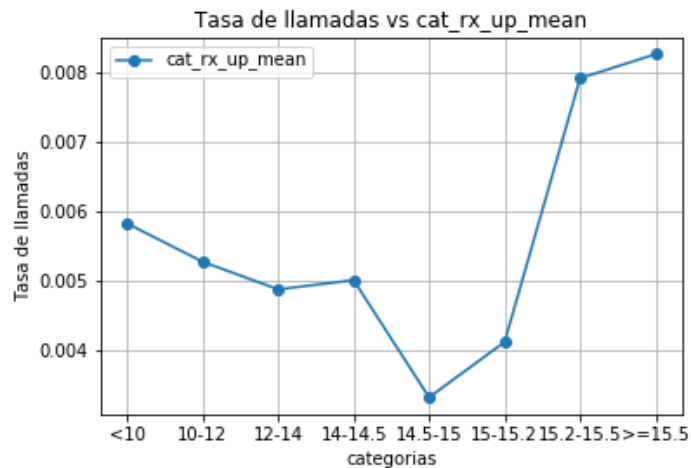


Figura 4.6: Tasa de llamados vs RX Power en Upstream

En la Tabla 4.5 se observa cada una de las 12 métricas de desempeño, el umbral determinado bajo las tasas de llamadas de forma empírica y el valor que indica el proveedor que son los umbrales operacionales. Los gráficos de cada una de las variables se encuentran en Anexo.

Se observa que varias de estas variables no tienen una recomendación y no se utilizan en la operación actualmente para ninguna tarea, estas variables están marcadas como NA para el umbral operacional.

Tabla 4.5: Umbrales operaciones vs umbrales propuestos en las métricas de desempeño de los cable módem

Variable	Umbral operacional	Umbral propuesto
<b>CER DN</b>	<1	0
<b>CER US</b>	<5	<5
<b>SNR DN</b>	>33	>30
<b>SNR US</b>	>30	>27
<b>CCER DN</b>	NA	<50
<b>CCER US</b>	NA	<50
<b>RX DN</b>	[-12,12]	[-10,10]
<b>RX US</b>	[14.5,15.5]	[14.5,15.2]
<b>TX US</b>	[25,55]	[25,51]
<b>T3</b>	NA	<2
<b>T4</b>	NA	0
<b>RESET</b>	NA	0

### 4.3.2. Preparación de los datos

Como se menciona anteriormente, se toma la llamada técnica de internet del día como la variable objetivo y para las métricas de desempeño de los cable módem de los clientes se realizan agrupaciones diarias, tales como: el mínimo, máximo, promedio y varianza del día, también se contabiliza la cantidad de horas con las variables en estado nulo y la cantidad de horas con la variable fuera del umbral operacional y umbral propuesto tomando este set de variables como variables predictoras.

Para la selección de variables se utiliza la prueba de Kolgomorov-Smirnov (KS), la cual nos entrega una medida de distancia entre las distribuciones de las variables para cada categoría (llama o no llama). Entonces, mientras mayor es el valor del índice ks la variable puede separar las categorías de mejor forma. Como ejemplo, en la Tabla 4.6 se comparan las variables que contabilizan los valores fuera de rangos operacionales y propuestos para seleccionar las que tienen un mayor índice ks.

Uno de los problemas del aprendizaje de esta tarea es el desbalanceo de las clases, ya que tan solo el 0.4% de las muestras cumplen con la condición de ser llamadas técnicas y el 99.6% de las muestras pertenecen a la categoría 0. Para solucionar este problema se utilizan 2 estrategias, la primera donde se extraen todas aquellas muestras que cumplan con el comportamiento de las llamadas y se realiza un muestreo estratificado del otro grupo respecto a las zonas que pertenecen y la antigüedad que tienen, tomando solo un porcentaje de los datos de este grupo y teniendo de esta forma el dataset de entrenamiento balanceado y con me-

Tabla 4.6: Ejemplo selección de variables índice ks

Variable	Umbral Operacional	Umbral Propuesto
<b>CER DN</b>	0,04	<b>0,08</b>
<b>CER US</b>	0,08	<b>0,08</b>
<b>SNR DN</b>	<b>0,05</b>	0,04
<b>SNR US</b>	0,05	<b>0,07</b>
<b>CCER DN</b>	NA	<b>0,01</b>
<b>CCER US</b>	NA	<b>0,03</b>
<b>RX DN</b>	0,01	<b>0,02</b>
<b>RX US</b>	0,07	<b>0,08</b>
<b>TX US</b>	<b>0,03</b>	0,02
<b>T3</b>	NA	<b>0,10</b>
<b>T4</b>	NA	<b>0,02</b>
<b>RESET</b>	NA	<b>0,05</b>

nos registros que la base original, a esta estrategia se le llama balanceo de clases por sub muestreo.

La segunda estrategia para afrontar este problema es tomar el dataset desbalanceado y otorgar pesos a las diferentes clases. El peso que se les va a otorgar a cada clase está dado por la implementación realizada por la librería de python sklearn y basada en una heurística de balanceo la cual le entrega un peso inversamente proporcional a la cantidad de muestras de la clase en el dataset, es decir que si el dataset está completamente balanceado el peso para cada clase es igual a 1. El peso utilizado es el siguiente:

$$peso_j = \frac{n_{filas}}{n_{clases} * n_{muestras_j}} \quad (4.1)$$

### 4.3.3. Modelamiento

Para el modelamiento se toman 28 días del mes de julio como data de entrenamiento, debido a que este mes tiene el peak de llamadas técnicas, por lo que hay más registros con el comportamiento que se quiere aprender. Entonces, con los 28 días se generan 2 dataset uno balanceado y otro desbalanceado al que se le aplicará pesos a las clases. Finalmente se utiliza 1 semana del mes de septiembre completa (sin balancear) como data de testeo de los modelos, para observar cómo se comporta el modelo en un período donde las llamadas son más estables.

Los modelos que se utilizan para buscar patrones en el comportamiento son: Regresión logística, Árbol de decisión, Random Forest y Xgboost. Estos algoritmos se entrenan en la base de entrenamiento por medio de k-fold validation. Al finali-

zar el entrenamiento, se compara el desempeño de cada algoritmo en la base de testeo teniendo los resultados que se muestran en la siguiente parte.

## 4.4. Evaluación de los modelos

Para evaluar los modelos se utiliza la métrica del área bajo la curva ROC (AUC), el recall que corresponde al porcentaje de muestras de la clase que se logran capturar y la precisión que corresponde al porcentaje de muestras que se clasifican de forma correcta. Considerando que la etiqueta de entrenamiento es una aproximación a una falla en el servicio de internet, el área bajo la curva es una mejor métrica para indicar qué algoritmo ordena de mejor forma a los clientes de acuerdo a la probabilidad de obtener un problema con el servicio.

Tabla 4.7: Evaluación algoritmos de clasificación

Modelo	AUC	Precision	Recall
Regresión logística (balanceado)	70.4	0.97 %	50 %
Regresión logística (pesos)	70.7	1.17 %	50 %
Árbol de decisión (balanceado)	79.2	1.77 %	50 %
Árbol de decisión (pesos)	78.8	1.66 %	50 %
Random forest (balanceado)	79.4	1.73 %	50 %
Random forest (pesos)	79	1.74 %	50 %
<b>Xgboost (balanceado)</b>	<b>80</b>	<b>2 %</b>	<b>50 %</b>
Xgboost (pesos)	79.8	1.85 %	50 %

En la Tabla 4.7 se observan los resultados de los 4 algoritmos que se entrenaron de las 2 formas: balanceando las clases por medio de un submuestreo y otorgándole pesos a las clases. Luego se utiliza un punto de corte dado que capture el 50 % de las llamadas, por lo que el resultado en el recall va a ser idéntico para todos los modelos. Dado este punto de corte, se busca qué algoritmo tiene un área bajo la curva más alta y logra una mayor precisión, siendo estos los criterios utilizados para elegir el modelo.

Como se observa en la Tabla 4.7 , el algoritmo que tiene mejor desempeño en las métricas de área bajo la curva ROC y de precisión es el algoritmo xgboost entrenado en la data balanceada, por lo que es el algoritmo escogido para ser utilizado en el proyecto y el proceso de visitas técnicas proactivas. Para tener una idea de comparación, el porcentaje de llamadas globales en la base de entrenamiento es de 0.4 % aproximadamente, entonces el modelo captura la mitad de las llamadas con un lift de 5 veces las llamadas de la cartera completa.

## 4.5. Resultados del modelo seleccionado

Se puede observar en el modelo seleccionado, la curva de ganancia en la Figura B.1, que muestra el eje y el porcentaje de llamadas que se capturan al tomar cierto porcentaje de los clientes dado por el eje  $x$ .

Para elegir algún punto de corte al problema, se utiliza la métrica del f1-score que es la media armónica entre la precisión del modelo y el recall. En la Figura 4.7, se muestra los diferentes resultados para cada punto de corte, donde se observa que el f1-score toma su máximo valor cuando el punto de corte es 0.8. También se puede observar en la línea azul que corresponde a la precisión del modelo, como éste logra ordenar a los clientes en grupos de mayor a menor tasa de llamados.

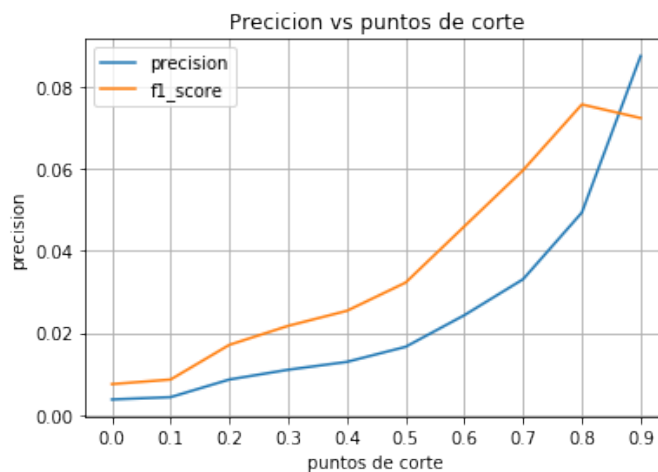


Figura 4.7: Precisión del modelo vs puntos de corte

Finalmente, se observa qué variables tienen mayor importancia para el modelo descrito anteriormente. En la Figura 4.8, se muestran las 10 variables con mayor importancia para el modelo. Dentro de éstas se observa que la primera variable tiene relación con algo que se descubrió en el análisis exploratorio, que muestra que los valores nulos de las variables no se dan únicamente por fallas en la carga, sino que muestran algún problema relacionado con la calidad del servicio. Dentro de las métricas técnicas, se ve que las que tienen mayor relevancia son los indicadores T3 que corresponde a contadores asociados a los protocolos de conexión del cable módem con las ondas portadoras del servicio, lo cual indica algún problema relativo a la comunicación con los sistemas que proveen el internet. Luego, vienen las variables de CER, tanto de bajada (downstream), como de subida (upstream), que corresponden al porcentaje de paquetes que llegan con error y no pueden ser corregidos. Finalmente, también se observa que la variación en la fuerza de recepción de la señal puede indicar problemas en el servicio del cliente.

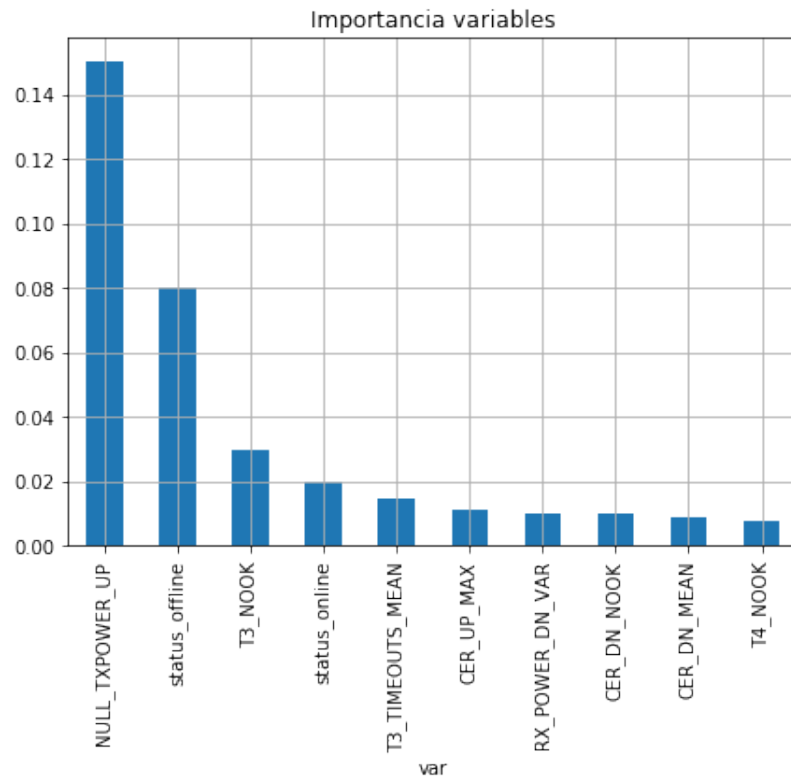


Figura 4.8: Importancia de las variables

# Capítulo 5

## Propuesta de Apoyo Tecnológico

En este capítulo se va a exponer la arquitectura TI del proyecto y se muestre el flujo de las herramientas de soporte tecnológico en el proceso. Principalmente, se expondrá cuáles son las plataformas y herramientas necesarias para la captura, transformación y carga de los datos, luego la forma en que se va a utilizar el modelo antes presentado. Finalmente la forma en que este se va a mostrar a los usuarios encargados de realizar las tareas del proceso.

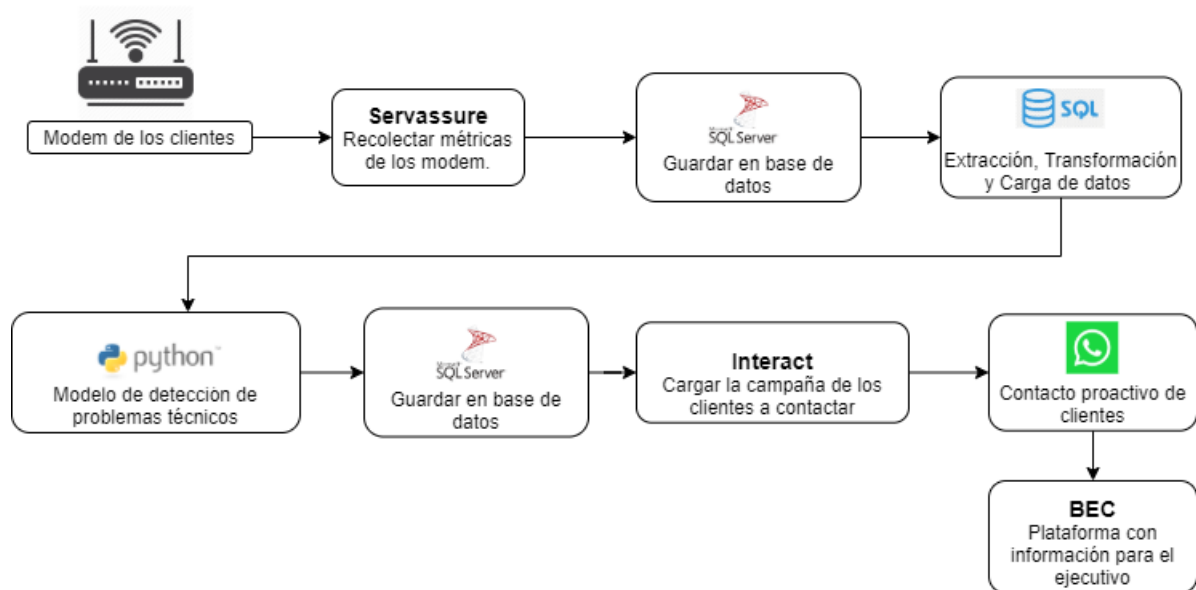


Figura 5.1: Propuesta de arquitectura TI del proyecto

En la Figura 5.1 se expone la arquitectura TI propuesta, la cual utiliza en su mayoría herramientas y plataformas, actualmente utilizadas por la organización. La principal razón de los anterior es adaptación de los usuario con estás herramientas.

El flujo comienza con la recolección de datos, que se realiza por medio de una plataforma llamada Servassure. Servassure tiene como función consultar las métricas de desempeño a de los clientes activos; la recolección de los datos se realiza

cada hora, pero la carga de éstos a las bases de datos se carga al final del día; es decir, una vez finalizado el día y cuando el uso de las bases se ve reducido comienza el procedimiento que carga los datos del día anterior. El motor de bases de datos que utiliza VTR es Microsoft SQL Server. Posterior a la recolección de datos viene la transformación de estos. Esta parte se realiza por medio de procedimientos almacenados escritos en lenguaje sql, donde el área BI automatiza un proceso ETL (Extracción, transformación y carga) que permita consultar las bases de datos y generar las variables que van a ser utilizadas para alimentar el modelo previamente entrenado. Una vez que se tenga las variables ya generadas, éstas van a ser consumidas por una API generada en python y almacenada en los servidores de la compañía, donde va a tener como input los datos del cable módem con las variables requeridas por el modelo y como output cada MAC de los clientes y la probabilidad de tener problemas con el servicio de internet fijo. La información resultante del modelo es cargada en una tabla alojada en las bases de sql Server.

Posteriormente, el área de clientes en conjunto con el área de operaciones deciden la cantidad de clientes a contactar de acuerdo a la capacidad de la operación y los priorizan de acuerdo a la probabilidad de tener un problema de internet, la cantidad de servicios que tienen contratados los clientes y el valor de los servicios. Con este dato el área de clientes carga de forma manual los clientes a ser contactados a una plataforma llamada Interact. Interact es la plataforma utilizada por VTR para realizar campañas a partir de los rut de clientes a contactar, el canal de contacto (correo electrónico, wsp, mensaje de texto o contacto telefónico), nombre de la campaña y mensaje que visualiza el ejecutivo tanto para las acciones proactivas, como para los ejecutivos del call center que atienda llamadas reactivas.

Por último, cuando el cliente recibe el contacto proactivo, por medio de wsp se despliega el siguiente mensaje: “Hola, desde VTR hemos detectado que tu servicio de internet fijo puede no estar funcionando con los estándares que requerimos. Por lo que te proponemos realizar una visita técnica gratuita, cuéntanos si deseas agendar una visita”. Un ejemplo de este mensaje se puede ver en la Figura 5.2





Figura 5.2: Contacto proactivo por mensaje de whatsapp

En caso que el cliente indique que desea agendar la visita, esta será coordinada por medio del canal junto al ejecutivo que tome el contacto. Para el caso contrario, el cliente estará marcado en una plataforma llamada BEC, plataforma que el ejecutivo utiliza cuando recibe llamadas del cliente y está le indica el estado comercial y técnico del cliente, como también las campañas que el cliente tiene cargada. Una vista de esta herramienta se encuentra en la Figura 5.3, 5.4 que muestran cuando el ejecutivo ingresa el rut del cliente y como se despliega una ventana con la campaña cargada indicando que el cliente es candidato a una visita proactiva y de esta manera agilizar su reclamo técnico.



Figura 5.3: Plataforma BEC - Ingresar rut del cliente

BEC - ONLINE

R.U.T. Cliente: 14726249-3      Dirección(es): NAPOLEON 3000, DEPTO 110 LAS CONDES      Nro Atención: 34874036

**Datos Cliente**

RUT: 0014726249-3      NOMBRE: JOSE VICTOR TERRAZAS TORRICO  
 COMUNA: LCON      DIRECCIÓN: NAPOLEON 3000, DEPTO 110  
 TELEFONO CONTACTO: 977711137, 977711137, 977711137      E-MAIL: jose.terrazas@vtr.cl, no tiene      NODO / CUADRANTE: 017 / 004

Ordnes Pendientes GOP: 0

**Campañas cargadas**

**Campaña Visitas Proactivas:**

- Cliente con problemas técnicos de internet en el domicilio.
- Ofrecer una visita técnica certificada sin costo.

**Servicios**

Servicio	Estado Comercial	Att BEC	Recl	Ord. Tec. BEC	Cont. BEC	Tot. Mes
Activo	Activo	6	0	0	0	0
Activo	Activo	0	0	0	0	0

CPE Asignados al Cliente y consulta on

Marca / Modelo: Arris - TG2492LG

Activar Administración WIFI:

Conectividad de equipo: Inalámbrico

Version Docsis: 3

Tipo de equipo: EMTA

Monto Cliente AVN:       Saldo: 20000      Limite: 20000     

Serie:Mac D-BOX: M91615EQE254

Customer ID: 0014726249-3\_2031264

SVA TV:

WikiBEC

Figura 5.4: Plataforma BEC - Campaña proactiva cargada

# Capítulo 6

## Evaluación del Proyecto

### 6.1. Evaluación Técnica

Con el propósito de efectuar la evaluación técnica, se hace un análisis retrospectivo del modelo en el cual se simula qué habría ocurrido si en cierto momento se utilizara el modelo para clasificar a los clientes con problemas críticos en el servicio de internet. Como VTR actualmente no ejecuta visitas proactivas, no será posible obtener el dato sobre el agendamiento de las visitas y cómo éstas impactan en la fuga de los clientes. Pero se puede observar, sobre la base de los clientes que tendrán en el futuro interacciones con VTR, sea por medio de llamadas o visitas técnicas.

Para realizar la simulación se utiliza la semana del 7 al 13 de Septiembre y se observa las llamadas y visitas técnicas reactivas del grupo durante 2 semanas; posterior a esto, también se observa el churn en un plazo más largo que corresponde a 90 días posterior a la semana de análisis.

La primera parte para la evaluación técnica consiste en utilizar el modelo para cada uno de los 7 días de la semana propuesta. El cual nos entregará una probabilidad de tener fallas del cliente para cada día de la semana y en base al punto de corte descrito anteriormente (0.8) se etiqueta si el cliente tiene problemas técnicos. Los resultados se observan en la Tabla 6.1, en ella se encuentran 62.000 clientes en los que se detectaron problemas en el servicio de internet entre el 7 y el 13 de septiembre y, por ende, 62.000 potenciales clientes para recibir las visitas técnicas proactivas. En estos se pueden observar algunas de las variables más relevantes del modelo, en todas hay un porcentaje mayor de cantidad de registros en umbrales deficientes. Tomando en cuenta que el grupo que se considera como falla de internet es un grupo crítico dentro de los clientes y moviendo el punto de corte a uno menor, se puede obtener más clientes que probablemente tengan problemas.

Para el grupo que se muestra anteriormente, se simulará qué pasa con aquellos que se ven con problemas técnicos según el modelo, estudiando las llamadas y

Tabla 6.1: Evaluación técnica - mirada retrospectiva

<b>Falla internet (según modelo)</b>	<b>Q</b>	<b>llama (07-13 sept)</b>	<b>NULL TXPOWER</b>	<b>CER_DN NOK</b>	<b>CER_UP NOK</b>	<b>T3 NOK</b>
no	1.1 M	2 %	1,8	7,5	4	7,2
si	62.000	11 %	19,5	14,5	15	14,6

visitas técnicas recibidas por esos clientes durante el período restante de septiembre y se observará qué porcentaje de esos clientes desconectaron el servicio de internet (fugaron) durante los 3 meses siguientes. Los resultados del análisis anterior, se muestran en la Tabla 6.2, donde el grupo con problemas técnicos tiene un mayor porcentaje de llamadas y visitas técnicas de internet que el grupo sano. Entonces, al simular un agendamiento al principio de la semana a los clientes del grupo de falla, se observa que al menos el 10 % aceptaría la visita, dado que ese porcentaje llamó por problemas técnicos en el sistema. A pesar de que los clientes que no llaman para quejarse, es probable que también tengan algún problema con el servicio, dado que como se observa en la Tabla 6.1 sus métricas de desempeño muestran una baja en la calidad. Por otro lado, se puede observar a los clientes que presentan problemas de internet, según el modelo un 7.9 % desconectó el servicio de internet en los meses futuros, lo que corresponde a un 41 % más respecto al resto de la cartera de clientes. Aunque el churn no estaba involucrado en el modelo éste tiene un impacto directo al ver los clientes con fallas, lo que comprueba la hipótesis inicial de la relación entre la calidad del servicio y la decisión de los clientes de desconectar el servicio de internet.

Tabla 6.2: Evaluación técnica - resultados

<b>Falla internet (según modelo)</b>	<b>Q</b>	<b>llama (14-30 sept)</b>	<b>Visita (14-30 sept)</b>	<b>churn (14 sept - 14 Dec)</b>
no	1.1M	4,2 %	1,5 %	5,6 %
si	62.000	10,8 %	3,9 %	7,9 %

De lo anterior, se desprende que por lo menos el 10 % de los clientes a contactar darían una respuesta positiva al tener un problema técnico en el servicio de internet. Además que el descrito, tendrá al menos un 41 % más de tasa de desconexión del servicio de internet a un plazo de 3 meses, para una acción en el grupo de corte entregado por el modelo como el corte que maximiza la métrica de F1 score.

Para otro análisis es importante observar la cantidad de TR que se ejecutan mensualmente en la operación de VTR. En la Figura 6.1 se observa la cantidad de TR efectuados mensualmente, donde se ve que en el momento peak de la pan-

demia se llegó a realizar aproximadamente 50 mil visitas técnicas al mes y que actualmente el número está estabilizado en cerca de 30 mil visitas. Esto significa que en el momento peak se realizaban cerca de 11 mil visitas de forma semanal y actualmente se realizan cerca de 7 mil visitas a la semana.

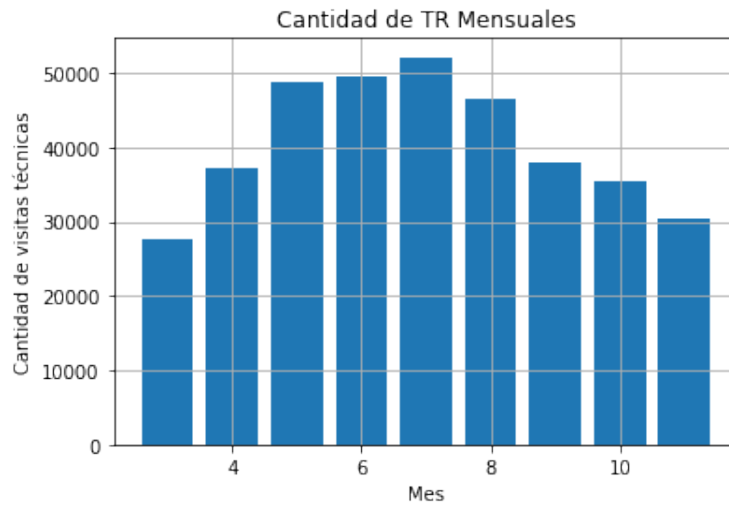


Figura 6.1: Cantidad de TR realizados mensualmente

Además, se sabe que hay un 10 % de capacidad no utilizada para las visitas técnicas hoy, lo que significa que de las 7mil visitas técnicas, es posible realizar unas 7.700 visitas técnicas sin problemas y llegar hasta 8.000. Es por esto, que se desea realizar un análisis retrospectivo que se adecue a estas cifras, es decir, que la cantidad de clientes que nos indican que tienen un problema técnico, en el futuro sea un número cercano a 1.000 clientes. Para obtener dichos números se corrige el puntaje de corte de 0.8 a 0.93 y se obtiene para la misma semana lo siguientes números:

Tabla 6.3: Evaluación técnica - mirada retrospectiva 2

Falla internet	Q	llama (07-13 sept)	NULL TXPOWER	CER_DN NOK	CER_UP NOK	T3 NOK
no	1,15 M	2,1 %	2,6	7,8	4,4	7,5
si	6.600	24 %	34	15,5	27	17

En la Tabla 6.3 se observa que la cantidad de clientes que quedan en este grupo crítico son 6.600 los cuales tienen una desviación mayor en cada una de las variables importantes.

En la tabla 6.3, se observa que la cantidad de clientes que quedan en este grupo crítico son 6.600 los cuales tienen una desviación mayor en cada una de las variables importantes.

Tabla 6.4: Evaluación técnica - resultados 2

Falla internet	Q	llama (14-30 sept)	Visita (14-30 sept)	churn (14 sept - 14 Dec)
no	1,15M	6,1 %	1,6 %	5,77 %
si	6.600	16 %	7,5 %	8,9 %

En la Tabla 6.4, se observa que del grupo de 6.600 clientes, el 16 % llama en las semanas posteriores del mes, esto deja un resultante de 1.000 clientes por lo menos que aceptarían una visita técnica con una tasa de fuga de 54 % más que el resto de la cartera. Cabe destacar que la aceptación de la visita técnica puede ser mayor a la mencionada en el análisis debido a que esto corresponde a una mirada retrospectiva donde no existe el proceso de visitas proactivas. Sin embargo, la regulación de los agendamientos se verá de forma semanal de acuerdo a las condiciones del sistema.

Finalmente, como dato relevante para la siguiente parte, los clientes que están en el grupo crítico de problemas técnicos tienen un 54 % más de desconexiones durante los siguientes 3 meses, pero la idea es que al realizar acciones sobre este grupo la tasa de desconexiones sea comparable al quintil con menor cantidad de problemas técnicos, por lo que el dato relevante viene dado por la Tabla 6.5, donde la diferencia porcentual de tasa de desconexión entre el grupo crítico y el quintil más sano es de un 93 %.

Tabla 6.5: Evaluación técnica - resultados grupo en buen estado

Falla internet	Q	llama (14-30 sept)	Visita (14-30 sept)	churn (14 sept - 14 Dec)
Quintil Sano	210.000	4,4 %	1,1 %	4,6 %
Medio	950.000	6,5 %	1,8 %	6 %
Crítico	6.600	16 %	7,5 %	8,9 %

## 6.2. Evaluación Económica

La evaluación económica requiere conocer los costos en que incurre el nuevo proceso y los beneficios esperados de éste, bajo ciertos supuestos y resultados de la evaluación técnica.

- Costos:** Los costos en que se incurre al implementar el nuevo proceso tienen 2 ítems: a) el costo de contactar a los clientes vía wsp, aproximado a \$500 pesos por cliente, y b) el costo de realizar las visitas técnicas, alrededor de \$15.000 por visita. Cabe destacar en este punto, que el costo de realizar una

visita técnica es despreciable para los primeros clientes, los que pueden ser cubiertos por la capacidad ociosa de técnicos regulares. El cálculo realizado corresponde a los primeros 750 clientes por semana aproximadamente, para mantener un poco de capacidad ociosa por posibles alzas, se contemplarán las primeras 700 visitas técnicas proactivas con la dotación de personal que actualmente existe. Como la meta es visitar un aproximado de 1.000 clientes semanales, el costo de los 300 clientes posteriores será nuevo para la organización.

- **Beneficios:** El principal beneficio corresponde a la disminución de churn del grupo de clientes con problemas técnicos. En la sección anterior se calculan las diferencias porcentuales, tomando el supuesto de que el grupo con alto churn tendrá una tasa de desconexiones del servicio de internet posterior a la intervención, parecida a los grupos sanos.

En la Tabla 6.6 se muestran los parámetros iniciales para observar el flujo de caja del proyecto.

Se observa que el churn a 3 meses del grupo es de 8.9% que corresponde a un 193% más que el grupo al cual queremos comparar. La cantidad de clientes semanales a contactar es de 6.000 y la tasa de aceptación suponemos que ronda el 16%. Luego, se observa que el churn al cual queremos llevar el grupo es de 4,61%. Los costos de contactar un cliente y de realizar una visita técnica son de \$5.00 y \$15.000 respectivamente y el valor del cliente por cada mes que permanece en VTR es de \$22.000. Finalmente, uno de los principales supuestos por sensibilizar es el menor porcentaje de clientes que se fugarán. Otro de los supuestos por sensibilizar, es la cantidad de meses de permanencia que tendrán los clientes que cambien su opción de desconectar el servicio de internet, que en una primera instancia es de 12 meses.

Tabla 6.6: Parámetros para la evaluación económica

Parámetro	Valor
churn	8,90 %
ganancia	193 %
# clientes semanales	6000
%aceptacion	16 %
%churn diff	4,61 %
Costo contacto	\$ 500
Costo Visita	\$ 15.000
Meses Permanencia	12
Valor cliente mes	\$ 22.000

Dado los parámetros que se muestran anteriormente, se calcula el flujo de caja mensual de un año como se muestra en la Tabla 6.7. Bajo el supuesto de que esto

se hará de forma recurrente cada semana y que las condiciones serán parecidas durante el año.

Tabla 6.7: Flujo de caja mensual

<b>MESES</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>...</b>	<b>12</b>
Clientes contactar mes	24000	24000	24000	...	24000
Clientes visitar mes	3720	3720	3720	...	3720
Ahorro en churn	160	160	160	...	160
<b>BENEFICIO</b>				...	
Ahorro churn	\$ -	\$ 42.117.493	\$ 42.117.493	...	\$ 42.117.493
<b>INGRESO</b>	\$ -	\$ 42.117.493	\$ 42.117.493	...	\$ 42.117.493
<b>Costos</b>				...	
Contacto	\$ 12.000.000	\$ 12.000.000	\$ 12.000.000	...	\$ 12.000.000
Visitas	\$ 13.800.000	\$ 13.800.000	\$ 13.800.000	...	\$ 13.800.000
<b>EGRESO</b>	\$ 25.800.000	\$ 25.800.000	\$ 25.800.000	...	\$ 25.800.000
<b>FLUJO CAJA</b>	-\$ 25.800.000	\$ 16.317.493	\$ 16.317.493	...	\$ 16.317.493

Dado el flujo de caja en la Tabla 6.7 se hace una copia para los 5 años posteriores, tomando nuevamente el supuesto de condiciones similares en los próximos 5 años, por lo que cada año el FC será similar al anterior. Finalmente se observa en la Tabla 6.8 los flujos de caja para los 5 años y se calcula el valor actual neto del proyecto con una tasa de descuento del 10 % obteniendo un VAN de \$582.615.206 pesos.

Tabla 6.8: Flujo de caja anual

<b>MESES</b>	<b>año 1</b>	<b>año 2</b>	<b>año 3</b>	<b>año 4</b>	<b>año 5</b>
Clientes contactar mes	288.000	288.000	...	...	288.000
Clientes visitar mes	44.640	44.640	...	...	44.640
Ahorro en churn	1.914	1.914	...	...	1.914
<b>BENEFICIO</b>					
Ahorro churn	\$ 463.292.424	\$ 463.292.424	...	...	\$ 463.292.424
<b>INGRESO</b>	\$ 463.292.424	\$ 463.292.424	...	...	\$ 463.292.424
<b>Costos</b>					
Contacto	\$ 144.000.000	\$ 144.000.000	...	...	\$ 144.000.000
Visitas	\$ 165.600.000	\$ 165.600.000	...	...	\$ 165.600.000
<b>EGRESO</b>	\$ 309.600.000	\$ 309.600.000	...	...	\$ 309.600.000
<b>FLUJO CAJA</b>	\$ 153.692.424	\$ 153.692.424	...	...	\$ 153.692.424

### 6.2.0.1. Análisis de sensibilidad

Para el flujo de caja expuesto en la Tabla 6.8 se realizará un análisis de sensibilidad que permita ver bajo qué supuesto el proyecto deja de ser rentable al plazo estu-



diado. El análisis de sensibilidad se va a realizar bajo 2 parámetros: el % de churn al cual vamos a llevar el grupo intervenido y la cantidad de meses de permanencia que tendrán estos clientes en promedio. El escenario que se vio se tomará como el escenario racional, ya que los supuestos fueron tomados del análisis retrospectivo que se realizó anteriormente y los parámetros que se van a utilizar serán los siguientes:

Tabla 6.9: Análisis de sensibilidad

VAN (10%)	MESES	<b>Pesimista</b>	<b>Racional</b>	<b>Optimista</b>
CHURN	-	8	12	15
<b>Pesimista</b>	170 %	-173M	327M	702M
<b>Racional</b>	193 %	-2,7M	582M	1.000M
<b>Optimista</b>	220 %	150M	814M	1.300M

En la Tabla 6.9 se muestra los diferentes escenarios: el churn en el escenario racional, se pretende llevar al nivel del grupo que está sano; en el escenario pesimista se supone que el churn será similar al grupo del medio que tenía problemas de red, pero no eran tan críticos como los del grupo intervenido y, finalmente, el escenario optimista, que dado que se realiza una medida de fidelización el churn es incluso menor que el grupo sano un 15 % más de diferencia porcentual que respecto a ese grupo, ya que serán clientes sanos a los que se les habrá realizado una medida de fidelización.

Respecto a los meses que se mantiene, se utiliza el supuesto de 3 periodos 8 meses en el escenario pesimista, 12 en el racional y 15 en el optimista. Dado este análisis de sensibilidad, es posible observar la relevancia de la permanencia de los clientes, ya que las acciones son más rentables si el valor del cliente en el tiempo es mayor (CLV), pero el cálculo de este se escapa del proyecto mencionado y queda como un trabajo futuro.

Respecto a los meses que se mantiene se utiliza el supuesto de 3 periodos 8 meses en el escenario pesimista, 12 en el racional y 15 en el optimista.

# Capítulo 7

## Conclusión y Trabajos Futuros

El proyecto nace de la contingencia del virus SARS-Cov-2 que provoca un cambio en la forma de vivir y relacionarse de las personas, debido a las cuarentenas impuestas por los gobiernos. Esta nueva forma de vivir exige mucho más a las compañías de telecomunicaciones de lo que se les había exigido hasta el momento. Por estas razones, los reclamos por parte de los clientes aumentan lo que golpea fuertemente a VTR. Este aumento en reclamos desencadena una mala imagen de la compañía de cara al cliente, provocando junto con el malestar del servicio un aumento mayor al 50 % en la tasa de fuga de los clientes de internet fijo.

La razón de las desconexiones se explica en gran parte por la mala calidad de servicio y la degradación en la experiencia de los clientes, sustentado además por encuestas de un porcentaje de clientes fugados y por bibliografía que muestra la relación entre estos fenómenos.

Por esta causa, se propone el diseño de un nuevo tipo de visita técnica que corresponde a la visita proactiva, es decir, se propondrá una visita técnica a aquellos clientes que se advierte tienen problemas en el servicio de internet antes que el cliente nos avise de este problema. Asegurando de esta forma que la experiencia del cliente no se vea degradada y utilizando la acción como una medida fidelizadora con los clientes.

A partir de un modelo de clasificación, se le asigna una probabilidad de falla a los clientes del servicio de internet fija, pudiendo llegar a los grupos de mayor riesgo considerando las variables técnicas e independiente a si el cliente inicia el reclamo o genera una llamada técnica debido a su problema. Sobre la base de este modelo, que permite validar la calidad del servicio entregado en términos técnicos, se monta el proceso de acciones proactivas. Cabe destacar que el modelo desarrollado durante el proyecto es útil no solo en las acciones proactivas, sino también en la priorización de acciones reactivas y como una forma de validar la calidad técnica del servicio. Sin embargo, son posibles usos que se escapan al alcance del proyecto.

Al realizar el análisis económico se puede observar que el proyecto es rentable bajo los supuestos realizados. Específicamente, es necesario considerar el porcentaje de churn que será ahorrado por el proyecto y el valor del cliente, en el presente nos indica qué tan rentable es la acción misma. Además de esto es importante destacar que se utiliza parte de la dotación de personal actual para efectuar las acciones proyectadas, ya que realizar una visita técnica tiene un costo alto.

Quedan propuestos varios desafíos en futuros trabajos orientados a mejorar el proceso de visitas proactivas; el primero, es medir el impacto del proceso propuesto en este trabajo y medir el impacto de aumentar la utilización de los recursos operacionales; el segundo, corresponde a un cálculo más preciso del valor de vida del cliente, de esta manera se pueden enfocar los esfuerzos en grupos de clientes más rentables; Por último, es posible realizar una segmentación de los problemas desde el punto de vista técnico, tomando como base el trabajo expuesto y algunos pilotos que ayuden a comprobar qué tipo de problemas corresponden las segmentaciones y cómo solucionarlos. Este último punto es importante para bajar el costo de las acciones proactivas o reactivas que se realicen, disminuyendo el tiempo y focalizando los esfuerzos de los técnicos. Disminuir los costos de las acciones es un desafío importante para el proyecto, con el fin de poder masificar más las acciones que reparen el servicio. De este modo los grupos menos críticos de clientes o con una menor diferencia porcentual en churn serían rentables para el proyecto.

# Bibliografía

- Barros, Ó. (2010). Ingeniería de negocios: Diseño integrado de negocios, procesos y aplicaciones ti cuarta parte. *Santiago*.
- Barros, Ó. (2015). Ingeniería de negocios: Diseño integrado de servicios, sus procesos y apoyo ti. *Santiago de Chile: Amazon, Kindle Edition*.
- Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.
- Dwivedi, Y. K., Papazafeiropoulou, A., Brinkman, W.-P., and Lal, B. (2010). Examining the influence of service quality and secondary influence on the behavioural intention to change internet service provider. *Information Systems Frontiers*, 12(2):207–217.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.
- Hax, A. C. and Wilde, D. L. (2003). El modelo delta-un nuevo marco estratégico. *Journal of Strategic Management Education*, pages 1–14.
- Johnson, M. W., Christensen, C. M., and Kagermann, H. (2008). Reinventing your business model. *Harvard business review*, 86(12):57–68.
- Mohri, M., Rostamizadeh, A., and Talwalkar, A. (2018). *Foundations of machine learning*. MIT press.
- Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.
- Smirnov, N. V. (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bull. Math. Univ. Moscou*, 2(2):3–14.
- Subtel (2019). Informe anual del sector de telecomunicaciones 2019.
- Subtel (2020). Informe trimestral del sector de telecomunicaciones.
- Suykens, J. A. and Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300.
- Svensén, M. and Bishop, C. M. (2007). Pattern recognition and machine learning.
- VTR (2018). Vtr reporte sostenibilidad.

Wirth, R. and Hipp, J. (2000). Crisp-dm: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, pages 29–39. Springer-Verlag London, UK.

# Anexo A

## Gráficos univariados de métricas vs tasa de llamados

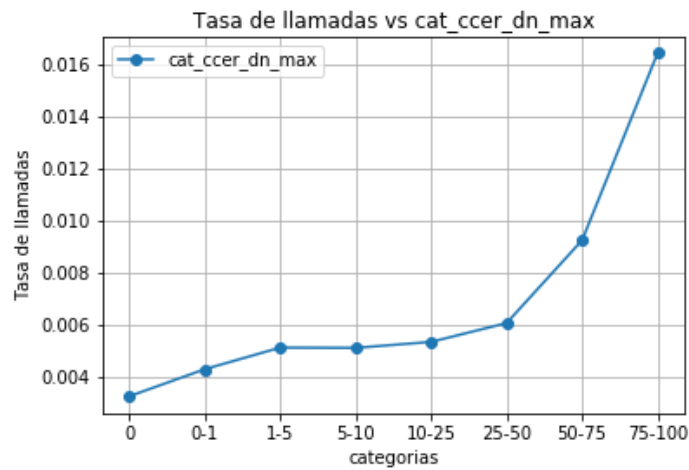


Figura A.1: Tasa de llamados vs CCER en Downstream

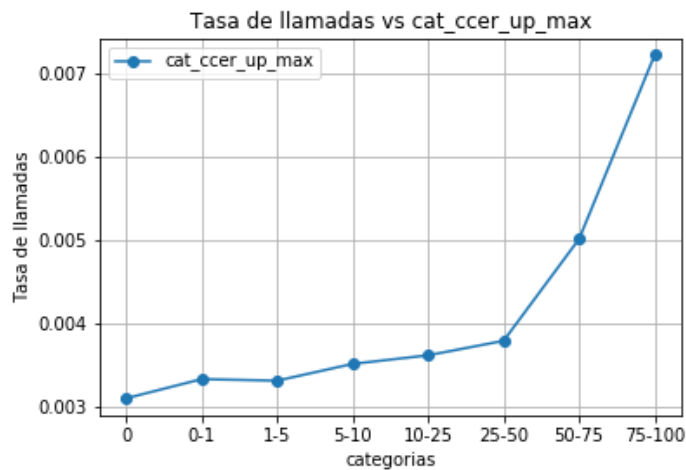


Figura A.2: Tasa de llamados vs CCER en Upstream

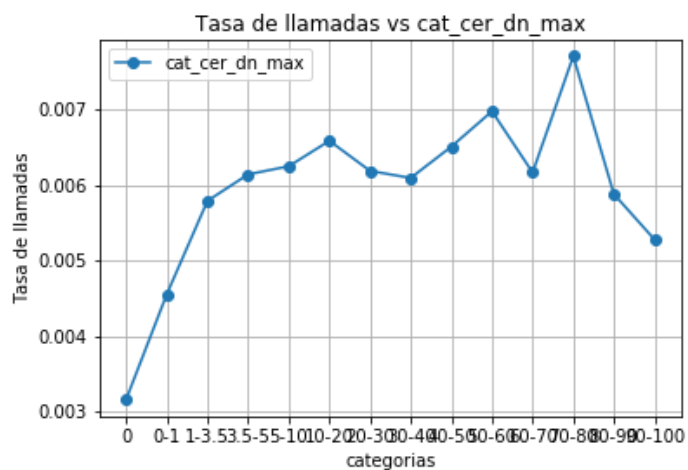


Figura A.3: Tasa de llamados vs CER en Downstream

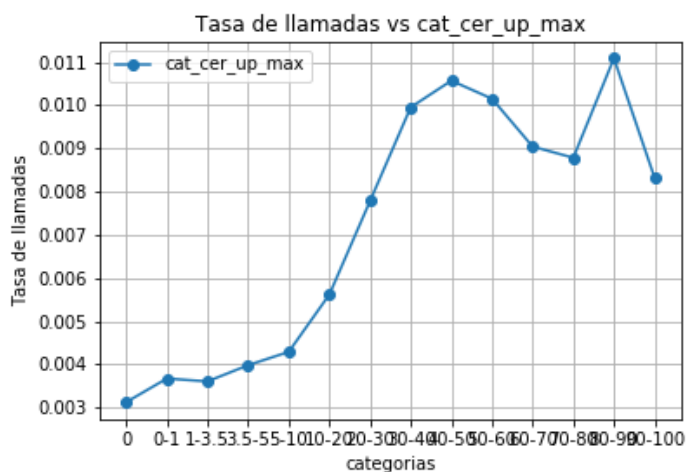


Figura A.4: Tasa de llamados vs CER en Upstream

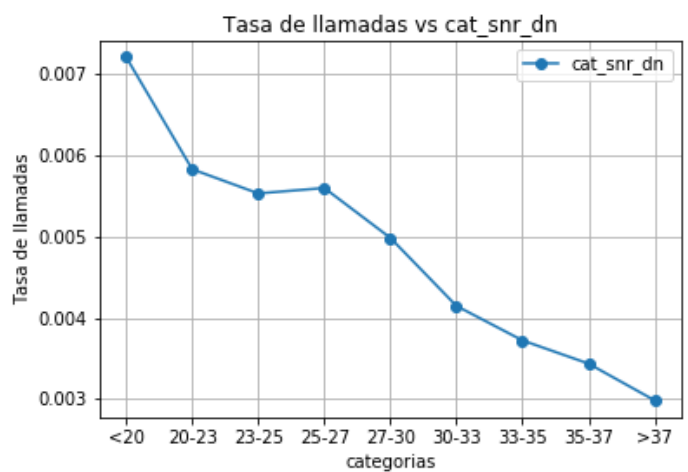


Figura A.5: Tasa de llamados vs SNR en Downstream

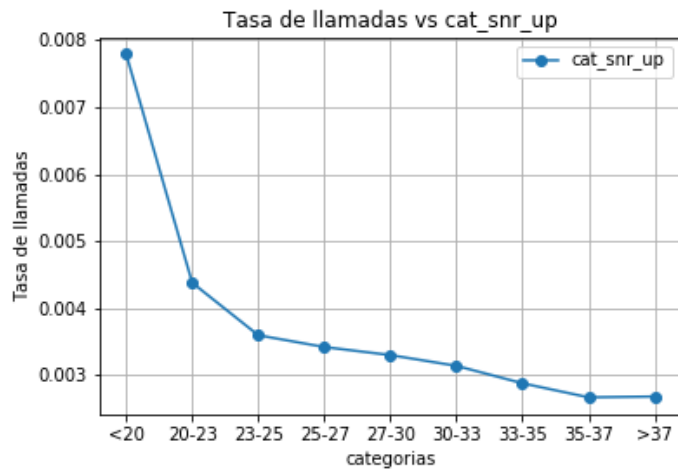


Figura A.6: Tasa de llamados vs SNR en Upstream

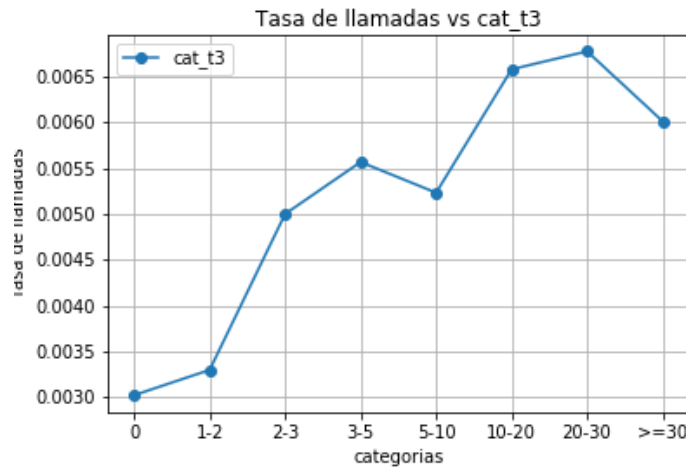


Figura A.7: Tasa de llamados vs T3 Timeouts

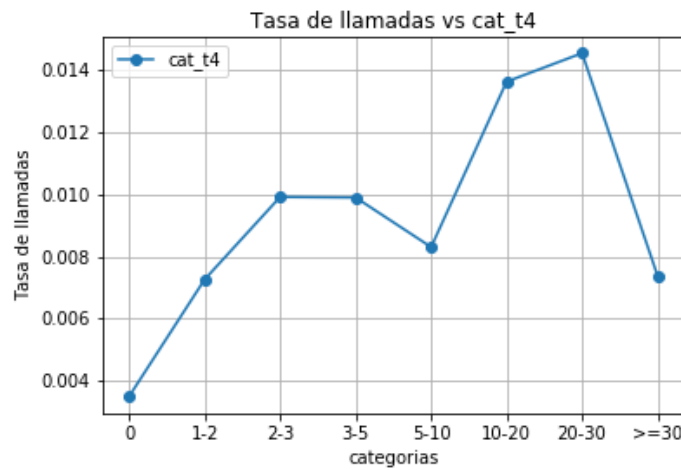


Figura A.8: Tasa de llamados vs T4 Timeouts



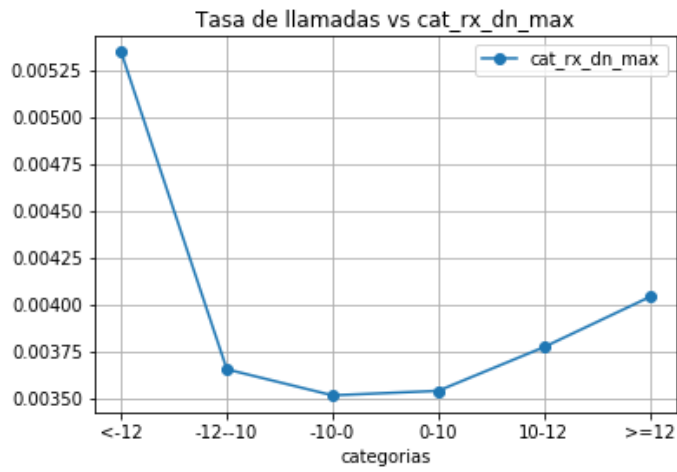


Figura A.9: Tasa de Llamados vs RX Power en Downstream

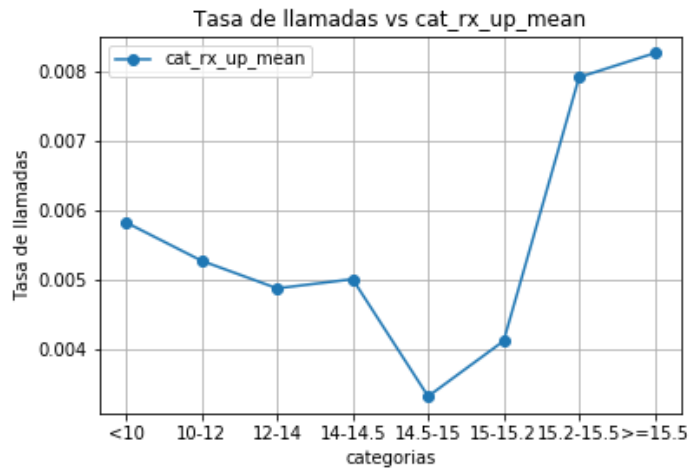


Figura A.10: Tasa de Llamados vs RX Power en Upstream

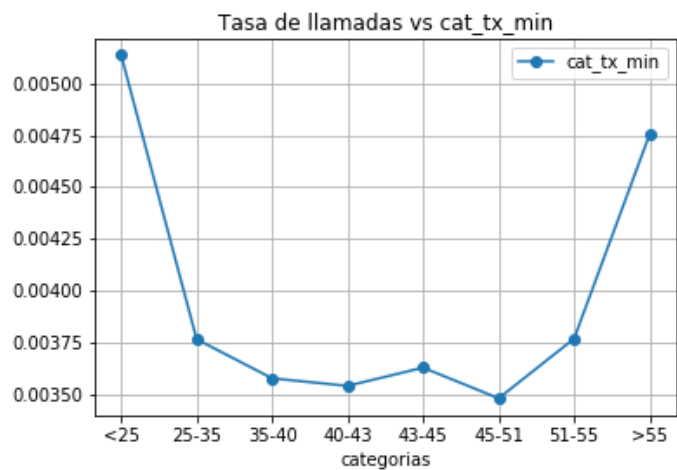


Figura A.11: Tasa de Llamados vs TX Power en Upstream

# Anexo B

## Resultados del modelo

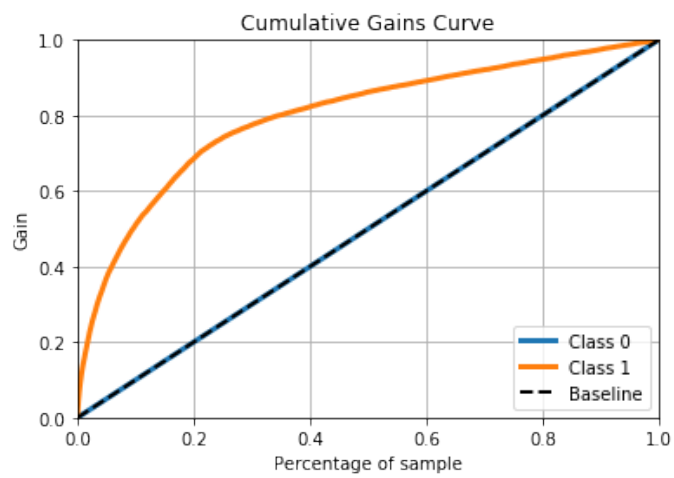


Figura B.1: Curva de ganancia del modelo seleccionado