



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

**REDISEÑO DEL PROCESO DE VENTA DE CRÉDITOS DE CONSUMO PARA CLIENTES
DEL BANCO RETAIL EN BCI**

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

OSCAR GABRIEL MARGUIROTT BRAVO

PROFESOR GUÍA:
RICHARD WEBER HAAS

MIEMBROS DE LA COMISIÓN:
OSCAR BARROS VERA
MARCEL GOIC FIGUEROA
FERNANDO LÓPEZ MOLINA

SANTIAGO DE CHILE
2016

RESUMEN EJECUTIVO

En la industria financiera, el objetivo de un banco es maximizar su utilidad. Ésta depende de tres factores: el precio, el volumen de venta y los costos (costo operacional y gasto en riesgo). A menudo los bancos asumen que la única variable de decisión para maximizar su utilidad es el volumen de venta, bajo un entorno competitivo donde no hay mucho espacio para competir por precio. Un enorme esfuerzo para reducir los costos (especialmente operacionales) también se ha implementado en los últimos años, pero con un impacto menor en las utilidades. Las estrategias de venta que adoptan los bancos asumen que el gasto en riesgo es un resultado de la operación del negocio, es decir, una consecuencia en lugar de una directriz. Esto ocasiona un enorme problema, que es que los bancos no conocen *ex ante* cuánto será el nivel de gasto en riesgo que se ocasionará producto de la venta de créditos, causando a veces pérdidas que no son capaces de predecir. Dos evidencias al respecto son que los precios a menudo no compensan los costos, causando pérdidas por la colocación de algunos créditos y ganancias por otros, y el sobre endeudamiento de los clientes (a través de ofertas más agresivas en términos de monto, es la principal causa del *default* del pago de créditos.

Desde 2014, Bci ha desarrollado un proyecto corporativo denominado *Plan de Transformación de Riesgo*, que consiste en posicionar el gasto en riesgo al mismo nivel que los ingresos por venta. Este plan adopta un nuevo paradigma, llamado *Apetito por Riesgo*, que permite que los bancos construyan sus estrategias considerando el gasto en riesgo como una directriz. Enmarcado en este plan, hay cuatro sub proyectos asociados que conforman el presente proyecto de tesis: rediseñar el proceso de generación de ofertas en campaña, desarrollar un modelo de predicción de gasto en riesgo, aplicarlo para establecer precios mínimos que aseguren *spread* libres de riesgo deseados y definir límites de endeudamiento óptimos. Estos proyectos fueron desarrollados utilizando técnicas del MBE y KDD, y son generalizables a otros problemas en otros tipos de organizaciones donde el gasto en riesgo puede ser utilizado como variable de decisión estratégica.

Estos proyectos fueron presentados y aprobados por el directorio de Bci, y están en ejecución desde marzo de 2015 como una estrategia *challenger* en un piloto controlado, con KPIs definidos y seguimiento periódico. De ser exitoso, podría incluso generarse oportunidades para reformular la política de créditos completa del Banco, y orientar la venta de créditos a cumplir objetivos de riesgo deseables.

AGRADECIMIENTOS

Al fin se cierra una de las etapas más importantes en mi vida, donde he madurado académica y profesionalmente. Esto no hubiera sido posible sin la ayuda y guía de todos mis compañeros, en especial a Max, por sus acaloradas discusiones. A Jaime, por sus sabios consejos. Felipe “manjarate” Vera, por siempre tener una paciencia para escuchar y la ocurrencia con que aconsejarme.

También quisiera aprovechar la ocasión de agradecerle a Fernando López, por creer en el proyecto y apoyarme para que se convirtiera en esta tesis. A mis compañeros de trabajo, por las acaloradas discusiones que tuvimos de las cuales nacieron los frutos de este trabajo.

Quisiera agradecerle también a Laurita y a Ana María por todo el apoyo y guía, y en especial toda la paciencia que tuvieron tanto conmigo como con todos mis compañeros de generación. A don Óscar Barros, por sus críticas que, aunque dolorosas a veces, fueron muy constructivas. A Richard Weber por su guía y orientación final, con su simplicidad maestra.

Finalmente quisiera terminar agradeciendo especialmente a todos los que estuvieron personalmente apoyándome estos últimos años. A mis padres, que aunque estando lejos, han sido un apoyo fundamental. A Sofía, que estuvo siempre a mi lado y yo la descuidé. Ahora desde el cielo me cuida y me inspira. A Rut, que ha tenido la suficiente paciencia para aguantar todas las noches en vela que pasé. ¡Eres lo máximo! Gracias también a Isa, mi chow chow, que cuando estaba bloqueado me invitaba a pasear y buscar la inspiración, técnica que siempre me funcionó.

Gracias MBE
Gracias compañeros de generación

Óscar Marguirott B.
Enero de 2016

A Sofía, siempre estuviste conmigo en este proceso y desde arriba me guiaste y apoyaste en los momentos más difíciles.

A Rut, mi fiel compañera. Gracias por tu paciencia y tu amor incondicional.

TABLA DE CONTENIDO

1	ANTECEDENTES GENERALES.....	1
1.1	LA ORGANIZACIÓN.....	1
1.2	MODELO DE NEGOCIOS	1
1.3	ESTRUCTURA ORGANIZACIONAL	3
1.4	CONTEXTO DE BCI EN LA INDUSTRIA BANCARIA	4
2	EL PROYECTO.....	8
2.1	ALGUNOS CONCEPTOS PREVIOS	8
2.1.1	<i>Riesgo de crédito.....</i>	<i>8</i>
2.1.2	<i>Límites de endeudamiento</i>	<i>9</i>
2.1.3	<i>Tipos de créditos de consumo</i>	<i>9</i>
2.2	EL PROBLEMA ACTUAL	9
2.3	ALCANCES Y OBJETIVOS DEL PROYECTO DE TESIS	10
2.4	CADENA DE VALOR	11
2.5	ACTIVIDADES DE LA CADENA DE VALOR.....	12
2.6	UNIDADES RELACIONADAS A LA CADENA DE VALOR	12
2.7	ANÁLISIS FODA.....	13
2.8	FUERZAS DE PORTER	13
2.9	MODELO DE NEGOCIO	14
2.10	PLANTEAMIENTO ESTRATÉGICO	15
2.11	IMPACTO DEL PROYECTO EN LA ESTRATEGIA ACTUAL	16
2.12	RESUMEN DEL PLANTEAMIENTO DEL PROYECTO	16
2.13	JUSTIFICACIÓN ECONÓMICA DEL PROYECTO	18
2.13.1	<i>Antecedentes para la evaluación</i>	<i>18</i>
2.13.2	<i>Construcción del Flujo de Caja.....</i>	<i>19</i>
2.14	GESTIÓN DEL CAMBIO DEL PROYECTO	24
3	ARQUITECTURA DE PROCESOS.....	26
3.1	PATRÓN DE NEGOCIOS.....	26
3.2	VARIABLES DE DISEÑO.....	27
3.3	ARQUITECTURA DE MACRO PROCESOS.....	28
3.4	DIAGRAMAS DE PISTAS	32
3.4.1	<i>Desarrollo de modelos de venta.....</i>	<i>32</i>
3.4.2	<i>Ejecución de modelos de venta.....</i>	<i>36</i>
4	CONSTRUCCIÓN DE LOS MODELOS DE VENTA.....	39
4.1	EL PROBLEMA.....	39
4.2	DEFINICIÓN DE LA VARIABLE OBJETIVO	39
4.3	DATOS DISPONIBLES	40
4.4	SELECCIÓN DE REGISTROS.....	41
4.5	SEGMENTACIÓN DE LA BASE	41
4.6	MODELAMIENTO.....	42
4.6.1	<i>Regresión lineal</i>	<i>42</i>
4.6.2	<i>Estimación por máxima verosimilitud.....</i>	<i>43</i>
4.6.3	<i>Redes neuronales.....</i>	<i>44</i>
4.6.4	<i>Support Vector Machines (SVM) y Support Vector Regression (SVR)</i>	<i>45</i>

4.6.5	<i>Construcción de una scorecard utilizando regresión logística</i>	46
4.7	TRANSFORMACIÓN DE LOS ATRIBUTOS	48
4.8	SOLUCIÓN DEL MODELO	49
4.8.1	<i>Segmento 1: clientes con antigüedad < 6 meses (clientes nuevos)</i>	49
4.8.2	<i>Segmento 2: clientes con antigüedad ≥ 6 meses y sin abono de remuneraciones en Bci (clientes sin abono de remuneraciones)</i>	62
4.8.3	<i>Segmento 3: clientes con antigüedad ≥ 6 meses y con abono de remuneraciones en Bci (clientes sin abono de remuneraciones)</i>	65
4.8.4	<i>Segmento 4: colaboradores</i>	68
4.9	APLICACIÓN DE LOS MODELOS PARA ESTABLECER UN ESQUEMA DE PRECIOS MÍNIMOS	69
4.10	APLICACIÓN DE LOS MODELOS PARA ESTABLECER LÍMITES DE ENDEUDAMIENTO ÓPTIMOS	70
4.11	PILOTO	72
4.11.1	<i>Diseño del piloto</i>	72
4.11.2	<i>Criterios de éxito del piloto</i>	73
4.11.3	<i>Aprobación y ejecución del piloto</i>	73
4.12	GENERALIZACIÓN	73
5	EL PROTOTIPO	76
5.1	DIAGRAMAS UML	76
5.1.1	<i>Diagrama de paquetes</i>	76
5.1.2	<i>Diagrama de casos de uso</i>	77
5.1.3	<i>Diagramas de secuencia</i>	77
5.2	TECNOLOGÍAS DE APOYO	78
5.3	ARQUITECTURA TECNOLÓGICA	79
6	CONCLUSIONES	81
7	BIBLIOGRAFÍA	83
8	ANEXOS	86
8.1	ANEXO 1: TABLA DE ENDEUDAMIENTO PROPUESTA SEGÚN MODELOS (SEGMENTOS 1 AL 3)	86

ÍNDICE DE TABLAS

TABLA 1: ANÁLISIS FODA	13
TABLA 2: INVERSIÓN INICIAL CONSIDERADA EN EL PROYECTO.....	20
TABLA 3: COSTOS ESPERADOS CON EL PROYECTO	21
TABLA 4: IINGRESOS ESPERADOS POR MEJOR ASIGNACIÓN DE LEADS EN CAMPAÑA	21
TABLA 5: INGRESOS ESPERADOS CON EL PROYECTO	22
TABLA 6: FLUJO DE CAJA, ESCENARIO CONSERVADOR (TODOS LOS VALORES EN M\$).....	22
TABLA 7: FLUJO DE CAJA, ESCENARIO MÁS PROBABLE	23
TABLA 8: FLUJO DE CAJA, ESCENARIO OPTIMISTA.....	23
TABLA 9: VARIABLES DE DISEÑO: ESTRUCTURA DE LA EMPRESA Y MERCADOS	27
TABLA 10: VARIABLES DE DISEÑO: COORDINACIÓN.....	27
TABLA 11: VARIABLES DE DISEÑO: PRÁCTICAS DE TRABAJO	27
TABLA 12: VARIABLES DE DISEÑO: INTEGRACIÓN DE PROCESOS CONEXOS.....	28
TABLA 13: VARIABLES DE DISEÑO: MANTENCIÓN CONSOLIDADA DE ESTADOS	28
TABLA 14: PARÁMETROS CONFIGURADOS PARA LA SIMULACIÓN DE DESARROLLO Y SEGUIMIENTO DE MODELOS DE VENTA.....	35
TABLA 15: RESULTADOS DE LA SIMULACIÓN DE DESARROLLO DE MODELOS DE VENTA.....	36
TABLA 16: PARÁMETROS CONFIGURADOS PARA LA SIMULACIÓN DE EJECUCIÓN DE MODELOS DE VENTA	37
TABLA 17: RESULTADOS DE LA SIMULACIÓN DE EJECUCIÓN DE MODELOS DE VENTA	38
TABLA 18: CUADRO DE PAGO DE EJEMPLO DE CRÉDITO QUE ENTRA EN CASTIGO	40
TABLA 19: VARIABLES CANDIDATAS PARA EL SEGMENTO 1	49
TABLA 20: RESULTADO DE LA REGRESIÓN LOGÍSTICA (CLIENTES NUEVOS).....	59
TABLA 21: PREDICCIÓN DE LA TG12 USANDO EL SCORE (CLIENTES NUEVOS).....	60
TABLA 22: RESULTADO DE LA REGRESIÓN LOGÍSTICA (CLIENTES SIN ABONO DE REMUNERACIONES).....	63
TABLA 23: PREDICCIÓN DE LA TG12 USANDO EL SCORE (CLIENTES SIN ABONO DE REMUNERACIONES).....	64
TABLA 24: RESULTADO DE LA REGRESIÓN LOGÍSTICA (CLIENTES CON ABONO DE REMUNERACIONES).....	66
TABLA 25: PREDICCIÓN DE LA TG12 USANDO EL SCORE (CLIENTES CON ABONO DE REMUNERACIONES).....	66
TABLA 26: EJEMPLO DE DETERMINACIÓN DEL LEVERAGE ÓPTIMO.....	71

ÍNDICE DE IMÁGENES

IMAGEN 1: CRITERIOS DE EXPERIENCIA BCI.....	2
IMAGEN 2: MODELO DE BALANCED SCORECARD DE BCI.....	2
IMAGEN 3: ORGANIZACIÓN INSTITUCIONAL DE BCI.....	3
IMAGEN 4: PRINCIPALES INDICADORES DEL SISTEMA BANCARIO EN CHILE AL AÑO 2012.....	5
IMAGEN 5: COLOCACIONES DE CONSUMO. COMPARACIÓN BCI VS SISTEMA FINANCIERO	6
IMAGEN 6: PROVISIONES SOBRE COLOCACIONES TOTALES. COMPARACIÓN BCI VS SISTEMA FINANCIERO	6
IMAGEN 7: ÍNDICE DE EFICIENCIA. COMPARACIÓN BCI VS SISTEMA FINANCIERO.....	7
IMAGEN 8: DIAGRAMA DE ACTIVIDADES Y PARTICIPANTES DEL PROCESO DE VENTA DE CRÉDITOS DE CONSUMO EN CAMPAÑAS MASIVAS.....	11
IMAGEN 9: POSICIONAMIENTO ESTRATÉGICO DE ACUERDO A MODELO DELTA	15
IMAGEN 10: CANVAS DEL PROYECTO	17
IMAGEN 11 PATRÓN DE ARQUITECTURA DE PROCESOS BP1	26
IMAGEN 12: ARQUITECTURA DE MACRO PROCESOS DEL PROYECTO	29
IMAGEN 13: DETALLE DE PROCESOS, MACRO 1	30
IMAGEN 14: ADMINISTRACIÓN DE LA RELACIÓN CON EL CLIENTE.....	31
IMAGEN 15: MARKETING Y ANÁLISIS DE CLIENTES.....	31
IMAGEN 16: GESTIÓN DE MODELOS DE VENTA BASADOS EN RIESGO.....	32
IMAGEN 17: ESQUEMA DE LA METODOLOGÍA KDD SEGUIDA EN EL DESARROLLO DE MODELOS DE BI.....	33
IMAGEN 18: DIAGRAMA DE PISTAS PARA DESARROLLO DE MODELOS DE VENTA	34
IMAGEN 19: DIAGRAMA DE PISTAS PARA EJECUCIÓN DE MODELOS DE VENTA.....	37
IMAGEN 20: SEGMENTACIÓN DE LA BASE	42
IMAGEN 21: EJEMPLO DE UNA REGRESIÓN LINEAL	43
IMAGEN 22: DEFINICIÓN CONCEPTUAL DE UNA RED NEURONAL	44
IMAGEN 23: DEFINICIÓN CONCEPTUAL DE LA TÉCNICA SVM	45
IMAGEN 24: EJEMPLO DE UNA REGRESIÓN LOGÍSTICA.....	46
IMAGEN 25: EJEMPLO DE TRAMIFICACIÓN DE LA VARIABLE EDAD.....	48
IMAGEN 26: HISTOGRAMA DE LA VARIABLE EDAD.....	50
IMAGEN 27: HISTOGRAMA DE LA VARIABLE RENTA LÍQUIDA ANTES DE LA TRANSFORMACIÓN	51
IMAGEN 28: HISTOGRAMA DE LA VARIABLE RENTA LÍQUIDA DESPUÉS DE LA TRANSFORMACIÓN	51
IMAGEN 29: HISTOGRAMA DE LA VARIABLE PROBABILIDAD DE INCUMPLIMIENTO (EN PORCENTAJES, DE 0 A 100).....	52
IMAGEN 30: HISTOGRAMA DE LA VARIABLE ANTIGÜEDAD EN CUENTA CORRIENTE.....	52
IMAGEN 31: HISTOGRAMA DE LA VARIABLE LEVERAGE PREVIO AL CURSE.....	53
IMAGEN 32: HISTOGRAMA DE LA VARIABLE LEVERAGE POST CURSE (CORREGIDO POR TRANSFORMACIÓN DE RENTA).....	54
IMAGEN 33: HISTOGRAMA DE LA VARIABLE NÚMERO DE ACREEDORES.....	54
IMAGEN 34: LÍNEAS DISPONIBLES SOBRE RENTA PREVIO AL CURSE.....	55
IMAGEN 35: HISTOGRAMA DE LA VARIABLE SCORE DE RIESGO (R04)	56
IMAGEN 36: DISTRIBUCIÓN DE LA VARIABLE TENENCIA DE CRÉDITOS HIPOTECARIOS	56
IMAGEN 37: ÁRBOL DE CLASIFICACIÓN PARA DETERMINAR LOS TRAMOS DE EDAD.....	57
IMAGEN 38: COMPARACIÓN GRÁFICA DE LA TG12 POR TRAMO DE SCORE (CLIENTES NUEVOS).....	61
IMAGEN 39: ESTABILIDAD DEL MODELO A TRAVÉS DEL TIEMPO (CLIENTES NUEVOS).....	61
IMAGEN 40: COMPARACIÓN GRÁFICA DE LA TG12 POR TRAMO DE SCORE (CLIENTES SIN ABONO DE REMUNERACIONES)	64

IMAGEN 41: ESTABILIDAD DEL MODELO A TRAVÉS DEL TIEMPO (CLIENTES SIN ABONO DE REMUNERACIONES).....	65
IMAGEN 42: COMPARACIÓN GRÁFICA DE LA TG12 POR TRAMO DE SCORE (CLIENTES CON ABONO DE REMUNERACIONES)	67
IMAGEN 43: ESTABILIDAD DEL MODELO A TRAVÉS DEL TIEMPO (CLIENTES CON ABONO DE REMUNERACIONES).....	67
IMAGEN 44: ÁRBOL DE CLASIFICACIÓN PARA EL MODELO COLABORADORES	68
IMAGEN 45: ESTABILIDAD DEL MODELO A TRAVÉS DEL TIEMPO (COLABORADORES).....	69
IMAGEN 46: DIAGRAMA DEL DOMINIO DEL PROBLEMA.....	74
IMAGEN 47: EJEMPLOS DE USO DE UNA SOLUCIÓN GENERALIZADA	75
IMAGEN 48: DIAGRAMA DE PAQUETES DEL SISTEMA.....	76
IMAGEN 49: DIAGRAMA DE CASOS DE USO DEL SISTEMA	77
IMAGEN 50: DIAGRAMA DE SECUENCIA PARA SOLICITA EJECUCIÓN DE MODELOS.....	78
IMAGEN 51: DIAGRAMA DE ARQUITECTURA TECNOLÓGICA SUGERIDA PARA EL PROTOTIPO...80	

1 ANTECEDENTES GENERALES

1.1 La organización

Bci se define como una “Corporación de Soluciones Financieras que participa en todos los negocios y operaciones financieras que la Ley General de Bancos le permite, ofreciendo a la comunidad productos y servicios con procesos de alta eficiencia operacional y excelencia en la calidad, con una permanente innovación tecnológica, prudentes políticas de administración de riesgos y exigentes estándares éticos, los que deben ser respetados por todas las personas que se desempeñan en sus empresas”¹. La misión de Bci asegura que se cumpla su visión, la cual es que “Aspiramos ser líderes regionales en innovación, cercanía y experiencia de clientes, y ser reconocidos como la mejor empresa para trabajar y desarrollarse”². Todo esto con el foco principal en el cliente, lo cual declara en su propósito común: “Hacemos que los sueños se hagan realidad a través de relaciones de confianza durante toda la vida de nuestros clientes”.

La estrategia de Bci, tal como la define la misma empresa, se basa en cuatro pilares básicos:

- *Los accionistas*, que tienen el poder de inversión y son dueños del capital del Banco
- *Clientes y proveedores*, que son primordiales para que la empresa se vuelva rentable
- *Colaboradores y sus familias*, que son el motor del Banco
- *La sociedad*, en el contexto que Bci se define como una empresa socialmente responsable

1.2 Modelo de negocios

Para llevar a cabo esta estrategia, se ha dispuesto un marco referencial consistente en cuatro criterios de experiencia que todos los colaboradores del Banco y sus empresas relacionadas deben cumplir, tanto de cara a sus clientes externos como también de sus clientes internos. Éstos se pueden visualizar en la Imagen 1.

A grandes rasgos, Bci evalúa la efectividad de sus estrategias mediante el uso de un modelo de *Balanced Score Card*³ corporativo, el cual se basa en 5 ámbitos de acción, tal como se muestra en la Imagen 2:

- **COLABORADORES:** la estrategia que define Bci en este ámbito es la de lograr atraer, desarrollar y retener talentos que desarrollen a la organización.

¹ Misión Bci. Fuente: memoria anual 2014 (Bci, 2014).

² Visión Bci. Fuente: memoria anual 2014 (Bci, 2014).

³ (Kaplan & Norton, 1996). La imagen piramidal del BSC de Bci ha sido definida por la propia organización.

SEGURIDAD	CERCANÍA	DILIGENCIA	IMAGEN
VELAR POR SU TRANQUILIDAD	CONOCER Y CONECTARSE EMOCIONALMENTE	ASESORAR Y EJECUTAR A LA PRIMERA	PROYECTAR LOS VALORES BCI EN CADA ACCIÓN
<ul style="list-style-type: none"> » Soy digno de confianza. » Doy continuidad al negocio. » Velo por el patrimonio del cliente y del banco. » Actúo con transparencia. » Velo para la seguridad de la información. 	<ul style="list-style-type: none"> » Yo conozco a mis clientes/colaboradores y demuestro interés por ellos. » Me pongo en el lugar del otro. » Me apropio de todos sus requerimientos. » Siempre demostrando la actitud de servicio de Bci. 	<ul style="list-style-type: none"> » Cumpló mis compromisos en tiempo y forma. » Tengo los conocimientos que necesito. » Simplifico la vida al cliente y lo mantengo informado en todo momento. » Mi trabajo es siempre impecable a la primera, sin errores y estando en el detalle. » Soy ágil. » Me anticipo. 	<ul style="list-style-type: none"> » Actúo en forma consistente con los valores: respeto, excelencia, integridad respeto y excelencia. » Cuido mi imagen personal y de mi equipo. » Genero un lugar de trabajo impecable. » Transmito que estoy orgulloso de pertenecer a la corporación Bci.

Imagen 1: Criterios de experiencia Bci
Fuente: memoria anual 2012



Imagen 2: Modelo de Balanced Scorecard de Bci
La representación piramidal corresponde a una definición interna de Bci
Confidencial

- **INNOVACIÓN:** Bci cuenta con una estrategia de generación de ideas que aporten al desarrollo organizacional, transversalmente en cualquier área de la corporación, orientadas siempre a mejorar la atención con los clientes.
- **PROCESOS:** contar con procesos ágiles, simples, libres de errores y siempre cerca del cliente es la estrategia de Bci referente al manejo de sus procesos internos.
- **CLIENTE:** para agregar valor al cliente, Bci sigue la estrategia de ser la fuente de otorgamiento de créditos más rápida y confiable, realizando una oferta de valor segmentada, innovadora y con la mejor experiencia de servicio multicanal.
- **FINANCIERO:** para los accionistas, la estrategia definida por Bci consiste en aumentar la utilidad con riesgo acotado sobre el capital.

El *Balanced Score Card* de Bci se lleva a la práctica a través de una serie de metas relacionadas a estos 5 ámbitos de acción, y se distribuye de forma personalizada a todos los colaboradores de la Corporación.

1.3 Estructura organizacional

Bci se compone actualmente de cuatro bancos, ordenados de acuerdo a la complejidad en el análisis de atención de sus clientes de menor a mayor (Imagen 3):

- *Banco Retail*, que atiende las necesidades financieras de crédito, ahorro y seguros de personas naturales.
- *Banco Empresarios*, que atiende las necesidades financieras de pequeñas y medianas empresas (PYME) cuyas ventas no superen las UF2.400 anuales.
- *Banco Comercial*, que ofrece asesoría comercial integral de productos y servicios financieros de acuerdo a las necesidades específicas de las empresas que tienen ventas por montos mayores a las UF2.400 anuales.
- *Corporate and Investment Banking (CIB)*, que tiene como fin proveer un modelo integral de atención para grandes corporaciones, clientes de alto patrimonio e instituciones financieras.

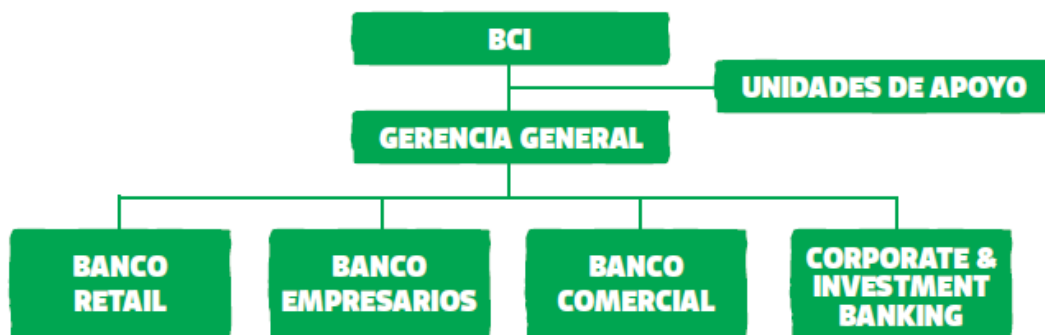


Imagen 3: Organización Institucional de Bci
Fuente: memoria anual 2014

Cada una de estas unidades se administra de manera independiente. Sin embargo, existen algunas gerencias al interior de estos bancos que jerárquicamente son transversales a dos o más de estos bancos.

Adicionalmente, existe una *Gerencia de Unidades de Apoyo* que presta servicios operacionales, tecnológicos, de gestión de personas y legales a los distintos bancos.

Los alcances de este proyecto se limitan al Banco Retail, el que, para efectos de simplicidad, denominaremos indistintamente Bci de aquí en adelante.

1.4 Contexto de Bci en la industria bancaria

Un banco que participa en el sistema financiero chileno tiene como principal fuente de ingresos la venta de créditos de consumo, y como principal fuente de costos el gasto en riesgo generado por estas colocaciones y los gastos operacionales propios del negocio. En lo referente a los ingresos, Bci se enmarca dentro de la industria bancaria en Chile como uno de los principales actores, con un 12.4% de participación sobre las colocaciones totales de consumo en el sistema financiero, indicador que, de acuerdo a la Imagen 5, ha venido al alza en los últimos 9 años. En lo referente a los costos, el gasto en riesgo de Bci en comparación al sistema financiero ha venido a la baja, situándose en diciembre de 2014 en torno al 2.13% de las colocaciones totales (ver Imagen 6), mientras que los gastos operacionales (medidos a través del índice de eficiencia) ha estado fuertemente descendiendo en los últimos 9 años, llegando a un 43.84% en 2014 (ver Imagen 7).

El sistema bancario chileno muestra un aumento del 37.9% de los activos totales entre 2009 y 2012, pero las utilidades netas han aumentado en tan sólo un 34.5% (ver Imagen 4), debido principalmente a que los costos del sistema financiero, sobre todo el costo operacional, han presentado un aumento (ver Imagen 7). Esto nos hace concluir que, para aumentar la utilidad de un banco, no basta con aumentar los ingresos por venta, sino más bien la estrategia se debe acercar a disminuir los costos operacionales y el gasto en riesgo. Para disminuir el primero, se necesitan procesos más ágiles, estandarizados y con poca intervención humana en las tareas más cruciales (por ejemplo, fomentando el uso de canales no presenciales para la venta, disminuyendo los tiempos de respuesta ante solicitudes de créditos, etc.). Sin embargo, en este proyecto nos centraremos en el segundo componente (gasto en riesgo). Para disminuir el gasto en riesgo hace falta colocar créditos a clientes con menor nivel de riesgo esperado y reducir la exposición en clientes con mayor nivel de riesgo, lo cual la industria ha venido haciendo con poca efectividad, debido al (casi) escaso uso de técnicas estadísticas en la predicción del nivel de gasto en riesgo de un cliente ofertado vía campaña.

SISTEMA BANCARIO
(MILES DE MILLONES DE PESOS)

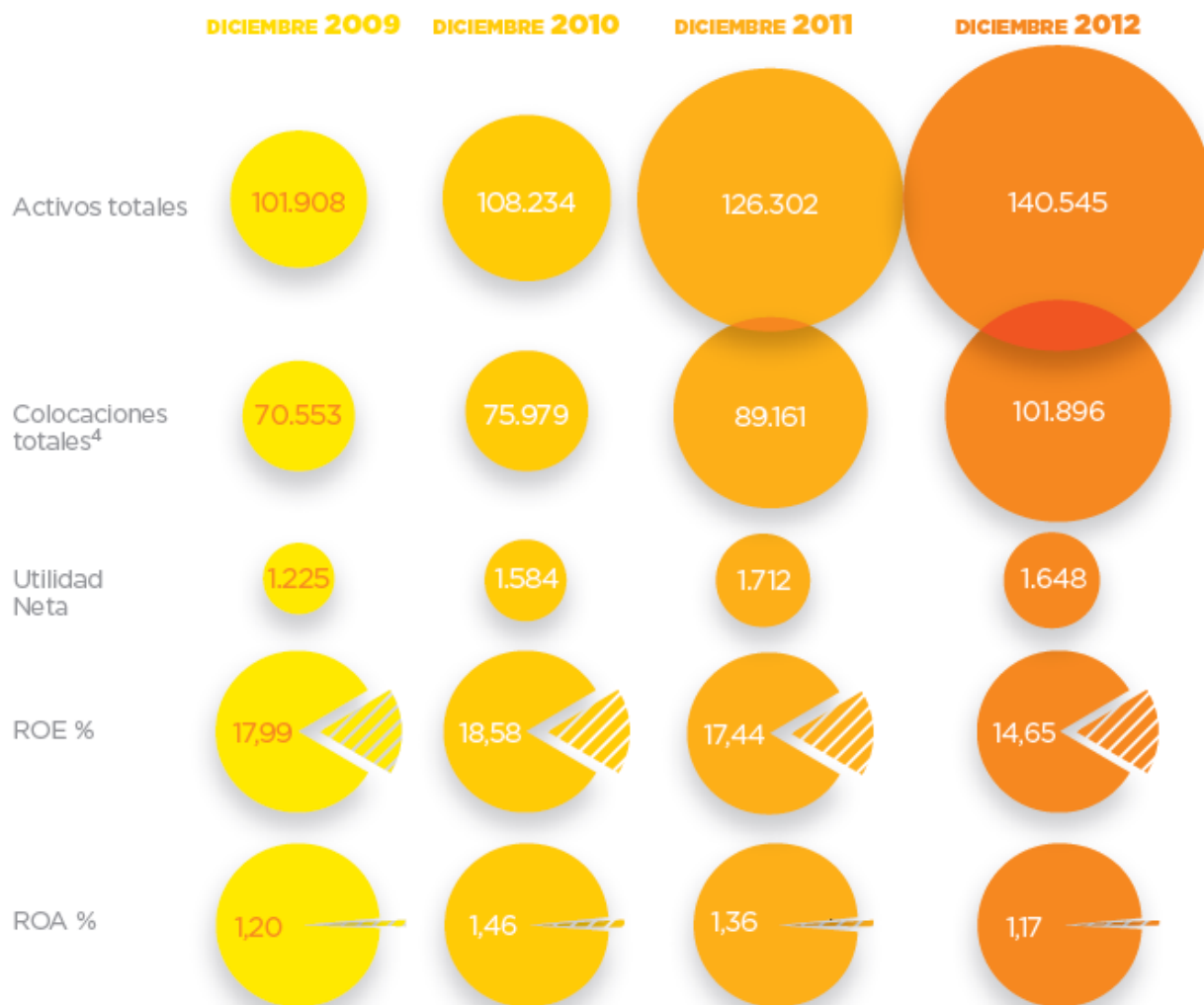
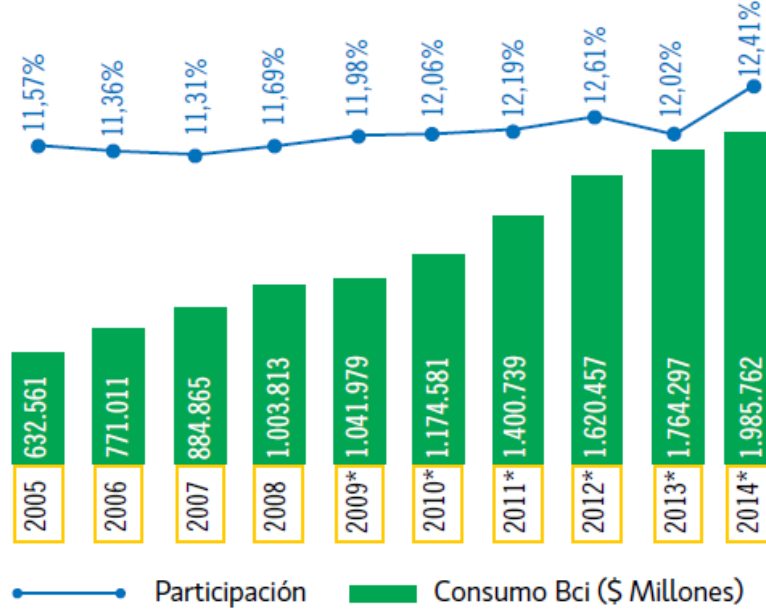


Imagen 4: Principales indicadores del sistema bancario en Chile al año 2012
Fuente: memoria anual Bci 2012 con datos de la SBIF

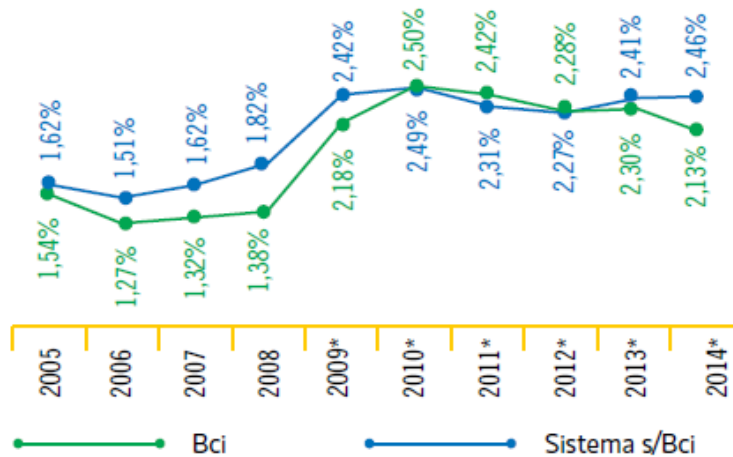
COLOCACIONES DE CONSUMO¹



(1) Participación de mercado sin considerar las adquisiciones de Corpbanca en Colombia.
 (*) Cifras bajo nueva normativa SBIF

Imagen 5: Colocaciones de consumo. Comparación Bci vs Sistema Financiero
 Fuente: memoria anual Bci 2014 con datos de la SBIF

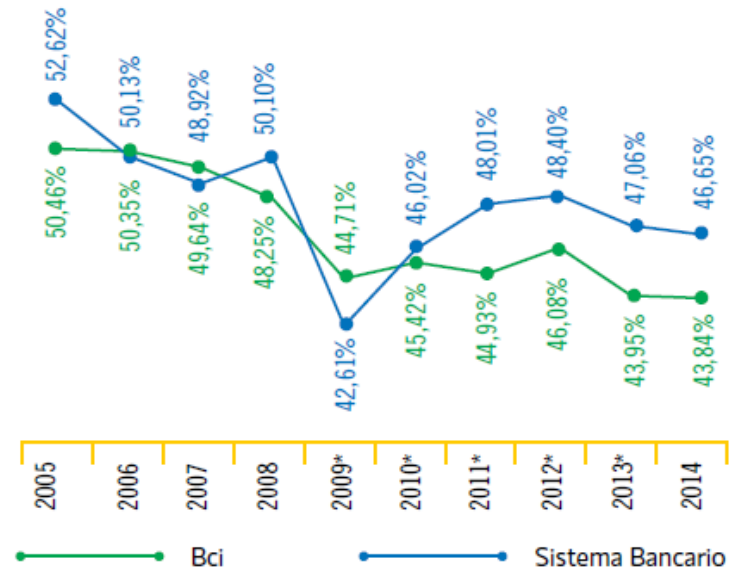
PROVISIONES SOBRE COLOCACIONES TOTALES: BCI Y SISTEMA BANCARIO¹



(1) Sistema Bancario excluye a Bci e incorpora inversión de Corpbanca en Colombia.
 (*) Cifras bajo nueva normativa SBIF

Imagen 6: Provisiones sobre colocaciones totales. Comparación Bci vs Sistema Financiero
 Fuente: memoria anual Bci 2014 con datos de la SBIF

ÍNDICE DE EFICIENCIA¹ BCI Y SISTEMA BANCARIO²



(1) Una baja de índice representa una mayor eficiencia

(2) Sistema Bancario excluye a Bci e incorpora inversión de Corpbanca en Colombia.

(*) Cifras bajo nueva normativa SBIF

Imagen 7: Índice de eficiencia. Comparación Bci vs Sistema Financiero

Fuente: memoria anual Bci 2014 con datos de la SBIF

2 EL PROYECTO

El presente proyecto de tesis consiste en la reformulación del proceso de generación de campañas de créditos de consumo para clientes de la Banca Retail. Antes de pasar a revisar el proyecto en sí, es necesario conocer algo de la nomenclatura que se maneja al interior de la organización.

2.1 Algunos conceptos previos

2.1.1 Riesgo de crédito

El riesgo de crédito corresponde al riesgo que los bancos y otras instituciones financieras que realizan préstamos corren, toda vez que existe una probabilidad de que los deudores dejen de cancelar sus obligaciones crediticias. El riesgo de crédito se calcula fundamentalmente con dos componentes: las provisiones por riesgo de crédito y la constitución de castigos.

Por norma, y de acuerdo a los criterios establecidos en los acuerdos de Basilea⁴, los bancos están obligados a provisionar un porcentaje de sus activos (créditos colocados en el sistema financiero) para que, en la eventualidad de que los créditos no se paguen, el patrimonio de los accionistas del banco se vea resguardado. Para esto, la *Superintendencia de Bancos e Instituciones Financieras* (SBIF) dispone de normas y procedimientos que permiten a los bancos calcular estas provisiones, las que generalmente son un porcentaje sobre el saldo de los créditos colocados. Estas provisiones, al ser dinero inmovilizado, le generan pérdidas a los bancos, que se contabilizan directamente en el estado de resultados.

Ahora bien, cuando los créditos que ha colocado un banco han superado los 180 días de impago (situación que se conoce como *castigo*), los bancos deben reconocer el 100% del saldo de ellos como pérdida. Estos se constituyen de manera directa en el estado de resultados del banco. Existe otro tipo de criterios para acelerar el castigo de créditos con menos de 180 días de impago, pero representan un porcentaje menor de la cartera castigada de Bci. Los créditos castigados no provisionan, puesto que se ha reconocido una pérdida total de ellos.

Ambos componentes se estiman de acuerdo a modelos estadísticos supervisados y validados por la SBIF. Sin embargo, estas provisiones y castigos se constituyen en función del *stock* de colocaciones (lo que el banco ya ha colocado en el sistema financiero), y no predicen

⁴ Los acuerdos de Basilea establecen criterios de administración del riesgo y el capital de instituciones financieras, a modo de recomendaciones, que diversos países han adoptado como forma de resguardar el patrimonio económico (público y privado) de las instituciones financieras.

necesariamente el comportamiento de los clientes en el pago de sus obligaciones de crédito para el *flujo* de colocaciones (es decir, la venta de nuevos créditos).

2.1.2 Límites de endeudamiento

El nivel máximo de exposición para un cliente que Bci está dispuesto a otorgar en operaciones de corto plazo se denomina *límite de endeudamiento* (o de crédito). Para su cálculo se consideran factores como los ingresos del cliente (su renta), indicadores de riesgo, deudas internas y externas, entre otros. Es responsabilidad de la Gerencia de Riesgo administrar los límites de endeudamiento en términos de montos, plazos, constitución de garantías asociadas, etc. Conceptualmente, el límite de endeudamiento se asocia al total de la deuda en el sistema financiero que adquiere el cliente, por lo que si a este límite se le descuenta la deuda actual de los clientes en todo el sistema financiero, se pueden calcular ofertas de créditos en campañas proactivas. Es necesario señalar que en el Banco Retail, las campañas masivas abarcan operaciones de corto plazo (créditos de consumo) dejando fuera operaciones de largo plazo (créditos comerciales e hipotecarios) para evaluación individual.

La definición de los límites de endeudamiento se enmarca dentro de la Política de Crédito del Banco Retail, que es administrada íntegramente por la Gerencia de Riesgo. Esta política está supeditada a la aprobación del directorio del Banco y es regulada por la SBIF.

2.1.3 Tipos de créditos de consumo

Existen dos tipos de créditos de consumo según el destino del crédito: créditos de consumo nuevos (que denominaremos *venta*) y créditos de consumo para refinanciamiento de deudas. Mientras el primero aumenta la exposición del cliente con el banco, el segundo tipo de créditos no aumenta la exposición, ya que sólo implica un ajuste en las condiciones originales pactadas de los créditos. Existen diversos motivos para refinanciar un crédito (rebaja de tasa, aumento de plazo, rebaja de cuota, situación de impago o renegociación, etc.), pero este proyecto de tesis se centrará principalmente en la venta de nuevos créditos de consumo, dejando fuera del ámbito de acción los refinanciamientos o renegociaciones de créditos.

2.2 El problema actual

Los bancos tienden a maximizar sus utilidades. En lo referente al negocio de colocar créditos de consumo, las utilidades dependen de tres factores principales: el precio cobrado por los créditos, el volumen de venta y los costos asociados. El precio es una variable que toma poca relevancia como variable de decisión en un mercado altamente competitivo, dejando como variables de acción para aumentar la utilidad el volumen de ventas y los costos. Generalmente, la estrategia que suelen seguir los bancos es la de incrementar las colocaciones para aumentar la utilidad, para lo cual se necesita destinar muchos recursos. Lo que se hace generalmente es competir por montos, entregando mayores ofertas que la competencia.

En lo referente a los costos, existen dos fuentes principales: costos operacionales (ej: mantención de sucursales, fuerza de venta, contratos asociados, etc.) y gasto por riesgo de crédito (provisiones y castigos). El paradigma actual es que el gasto en riesgo es una derivada del negocio bancario, no una variable de decisión. Por lo tanto, la estrategia que siguen los bancos al respecto es la de disminuir los costos operacionales al automatizar y estandarizar procesos, hacer un uso más eficiente de recursos, incentivar el uso de canales no presenciales (más baratos), etc.

Dicho lo anterior, en la banca chilena la generalidad es que los bancos establezcan modelos de venta basados en los ingresos, dejando el gasto en riesgo como una derivada que necesariamente se tiene que integrar en aquellos modelos. Como consecuencia de ello, la venta no necesariamente tiende a minimizar el gasto en riesgo, llevando a que los bancos no puedan controlar este gasto inicialmente. Como evidencia, se tiene que a veces un banco cobrando el mismo precio por el mismo producto a dos clientes distintos puede estar perdiendo utilidad en uno de ellos y ganándola en el otro. No hay una diferenciación por nivel de gasto en riesgo entre clientes.

El nuevo paradigma que está surgiendo recientemente en la industria bancaria es que los bancos tienen que definir su *Apetito por Riesgo*, que consiste en que el riesgo no sea la causa de la operación del negocio bancario, sino que el riesgo sea la directriz de la misma. La idea principal subyacente es que el banco debe establecer el nivel de riesgo que está dispuesto a asumir, y en base a ello se deriva el resto de iniciativas, incluidas la venta.

Adhiriendo a este concepto, Bci ha diseñado un plan corporativo de Transformación de Riesgo que tiene como fin, entre otros, situar la cultura de riesgo al interior de la organización, estableciendo lineamientos que colocan el gasto en riesgo como directriz de casi toda la actividad bancaria. Enmarcado dentro de este plan corporativo, hay tres sub planes que son tratados en este proyecto de tesis. El primero de ellos consiste en desarrollar modelos que predigan el nivel de gasto en riesgo para nuevos créditos otorgados, de manera de conocer *ex ante* cuál es la pérdida esperada de cada colocación. Relacionado a ello, el segundo sub plan consiste en aplicar este modelo para definir precios mínimos a cobrar, de manera de asegurar una rentabilidad mínima deseada individualizada para cada cliente. Finalmente, el tercer sub plan consiste en establecer el nivel de endeudamiento óptimo para no sobre endeudar a los clientes y no deteriorar su capacidad de pago.

El principal desafío consiste en rediseñar los procesos de generación de ofertas vía campañas masivas para aplicar estos modelos.

2.3 Alcances y objetivos del proyecto de tesis

El objetivo principal de esta tesis consiste en rediseñar el proceso de generación de campañas de créditos de consumo masivas. Dentro de este se establecerá la actividad de desarrollo de

modelos de venta basados en riesgo, enfocando los esfuerzos en colocar créditos vía campañas masivas de forma eficiente. Quedará fuera del ámbito de esta tesis definir el apetito por riesgo, rediseñar otras actividades de la cadena de valor distintas a la generación de campañas masivas y la modificación a la política de crédito.

2.4 Cadena de valor

La cadena de valor asociada al proyecto de tesis es la de venta de créditos de consumo nuevos a clientes de la banca de personas a través de campañas masivas. Esta cadena de valor se ejecuta mensualmente y representa alrededor del 85%⁵ de la venta de créditos de consumo nuevos. En la Imagen 8 se puede apreciar el diagrama de actividades relacionadas a este proceso, y los participantes que intervienen en cada una de las etapas. Si bien en este diagrama se describe el proceso completo, el ámbito del proyecto de tesis se centrará sólo en la primera actividad de la cadena de valor, que consiste en la generación de ofertas proactivas de créditos.

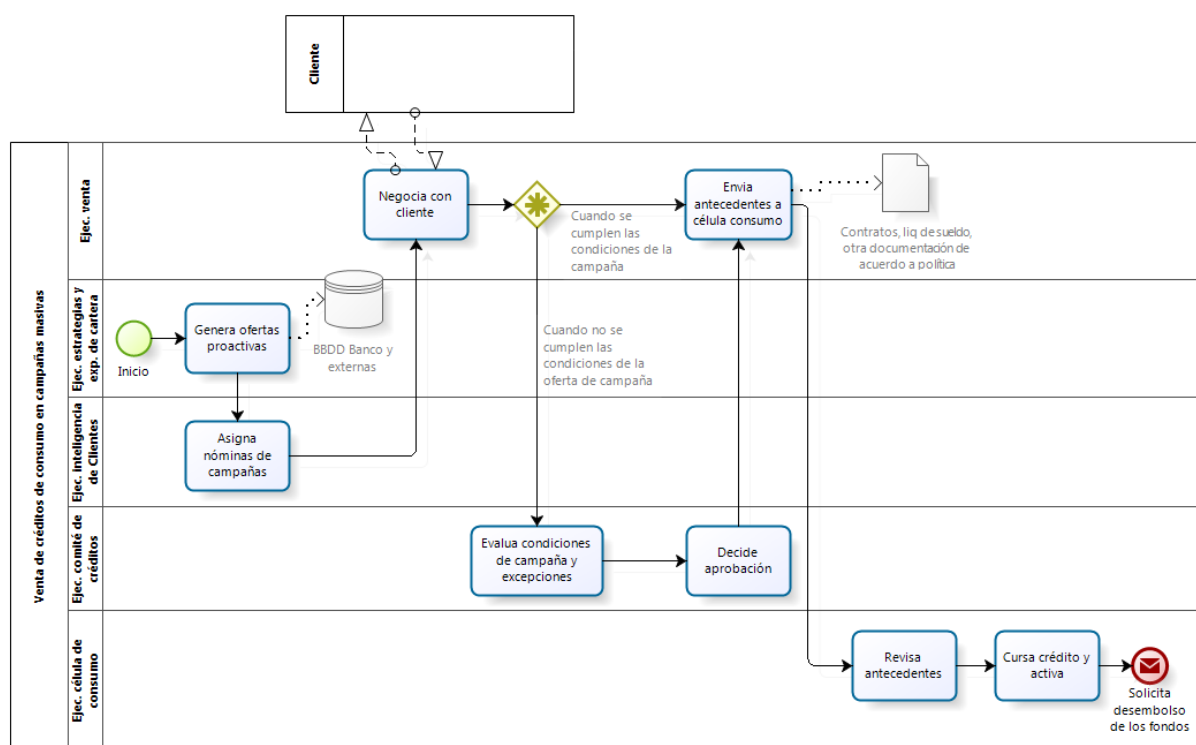


Imagen 8: diagrama de actividades y participantes del proceso de venta de créditos de consumo en campañas masivas

⁵ Fuente: datos internos Bci a abril 2015.

2.5 Actividades de la cadena de valor

- *Generación de ofertas proactivas*, donde se evalúan masivamente los clientes de acuerdo a la política de créditos para generar ofertas de campañas. Esta es la actividad dentro de todo el proceso donde se centrará el proyecto de tesis.
- *Negociación*, en donde se contacta a los clientes y se acuerda con ellos sus necesidades de financiamiento y las condiciones bajo las que se otorgaría eventualmente el crédito.
- *Evaluación*, tarea en la cual se aplica la política de créditos para decidir si el crédito es otorgado y bajo qué condiciones (en el caso que así sea requerido, por ejemplo, cuando un cliente solicita un monto diferente a lo ofrecido en la campaña o no se cumple alguna de las condiciones excepcionables de la misma).
- *Revisión de antecedentes*, actividad en la cual se revisa si el cliente cumple con los requisitos establecidos en la normativa para operar con el Banco y se validan las condiciones de curse.
- *Curse y activación*, donde se generan y firman los contratos, se genera la orden de desembolso de recursos y se procede a la activación contable de los créditos.

2.6 Unidades relacionadas a la cadena de valor

- *Sub gerencia de Explotación de Cartera y Estrategias*, dependiente de la Gerencia de Riesgo Retail, cuya función es la de generar las campañas de créditos y detectar nuevas oportunidades de negocios.
- *Unidad de Inteligencia de Clientes*, dependiente de la Gerencia de Marketing, que realiza la asignación de las nóminas de campaña a los distintos canales de venta y lleva el seguimiento de los indicadores claves de desempeño (KPI) de una campaña.
- *Canales de venta*, como sucursales, *telemarketing*, web, entre otros, que, dependiendo de o trabajando en conjunto a la Gerencia Comercial, ejecutan las campañas en términos de contactar a clientes, acordar negocios, recopilar antecedentes, etc.
- *Comités de Créditos*, unidad dependiente de la Gerencia Corporativa de Riesgo, que tiene como función principal la de evaluar la política de créditos en la originación misma, decidiendo aceptar o rechazar las solicitudes de créditos, siempre ayudando a configurar el mejor negocio entre el cliente y el Banco.
- *Célula de Originación*, unidad dependiente de la Gerencia de Operaciones, que tiene como función la revisión de la documentación relacionada con el proceso de venta, así como también la de activar los créditos cursados de acuerdo a la normativa interna y legal vigentes. Esta unidad se divide de acuerdo a la familia de productos que se valida y documenta. Existe la *Célula de Originación de Planes* (un plan se define como el *pack* de cuenta corriente + línea de sobregiro + tarjeta de crédito), la *Célula de Originación de Consumo* (que se orienta a los créditos de consumo), la *Célula Hipotecaria*, etc. Existe también una célula dedicada exclusivamente a la activación de forma transversal. Esta unidad se denomina *Célula de Activación y Curse* (o ACA).

2.7 Análisis FODA

Ante un escenario complejo y competitivo como la banca chilena, Bci ha adoptado distintas acciones para lograr el éxito en el sistema financiero. Estas acciones han aprovechado las oportunidades externas y las fortalezas internas, y en cierta medida también consideran las amenazas y debilidades. En la Tabla 1 se presenta brevemente las distintas características, tanto internas como de la industria, que competen a Bci y se relacionan con el actual proyecto.

Características internas	Características externas
FORTALEZAS <ul style="list-style-type: none"> ▪ Pionero en cultura de riesgo y apetito por riesgo ▪ Imagen corporativa sólida, cultura Bci, medidas de transparencia, educación financiera, etc. 	OPORTUNIDADES <ul style="list-style-type: none"> ▪ Precios no compensan los costos ▪ Los bancos aparentemente sobre endeudan a sus clientes.
DEBILIDADES <ul style="list-style-type: none"> ▪ Procesos de campaña y venta reactivos ▪ No hay explotación de nichos potenciales 	AMENAZAS <ul style="list-style-type: none"> ▪ Mercado de tarjetas de crédito creciente (producto sustituto)

Tabla 1: Análisis FODA

2.8 Fuerzas de Porter

- **Poder de negociación de los clientes:** el cliente tiene un mayor conocimiento de los productos bancarios debido a las medidas de transparencia de la industria. El cliente tiene más espacio para negociar. En el contexto actual, el poder de negociación de los clientes es una fuerza poderosa que hay que tener en cuenta al diseñar cualquier estrategia.
- **Poder de negociación de los proveedores:** se vuelve irrelevante en la industria bancaria (por lo menos en lo que a colocación de créditos respecta).
- **Amenaza de nuevos competidores:** si bien pueden aparecer nuevos competidores en la industria, es prácticamente imposible que alcancen un nivel de madurez suficiente en un corto periodo de tiempo. Los bancos de casas comerciales se están volviendo fuertes competidores en la industria, al competir bajo condiciones distintas que el resto de las instituciones financieras (comparten costos y clientes con el negocio comercial *retail*), pero su negocio se enfoca principalmente en el de colocación de tarjetas de crédito. Es una fuerza poco relevante, pero que en el mediano-largo plazo se podría volver una fuerza importante a considerar.
- **Amenaza de productos sustitutos:** como servicio, los sustitutos vienen a ser los créditos de consumo y tarjetas de crédito de otros bancos. El nivel de competencia aquí viene dado principalmente por precio. Adicionalmente, el financiamiento que entregan instituciones como las cooperativas de crédito o cajas de compensación (entre otras) también puede considerarse como un servicio sustituto al entregado por los bancos, pero

generalmente tienen precios mayores (debido a que sus costos son también mayores), por lo que actualmente no son tan relevantes en el análisis.

- **Rivalidad entre competidores:** hay una rivalidad creciente entre los principales bancos de la industria, y la pelea viene principalmente enfocada en calidad. Bci tiene un esquema de precios relativamente similar a la competencia, pero destaca respecto de los otros bancos debido a que su enfoque es en el servicio y la “cultura Bci”. Sin embargo, esta diferenciación es imitable en el tiempo, por lo que se debe tener en cuenta esta fuerza en el diseño de la estrategia de mediano y largo plazo de Bci.

2.9 Modelo de negocio

El modelo de negocio definido de acuerdo a Johnson⁶, debe resumir una metodología de reinención basada en tres etapas. Para que sea atractivo desde el punto de vista de los *stakeholders* debe considerar: primero, identificar cuál o cuáles son las oportunidades para satisfacer las necesidades de los clientes de una empresa; segundo, desarrollar un programa con cuatro elementos básicos para satisfacer dicha necesidad: propuesta de valor para el cliente, fórmula de utilidades, recursos clave y procesos clave; y tercero, comparar este nuevo modelo de negocios con el modelo de negocios actual de la compañía.

La oportunidad que se identifica es la de hacer más eficiente la venta de créditos de consumo en campañas, en conjunto con establecer precios mínimos de acuerdo al riesgo y límites de endeudamiento óptimos. La hipótesis que hay por detrás es que el riesgo de crédito es distinto para cada cliente, por lo que se busca predecir este gasto para cada cliente, y luego establecer un precio que asegure una ganancia en términos de *spread*⁷ para hacer el negocio rentable. En conjunto con ello, los límites de endeudamiento deben asegurar un gasto en riesgo deseado, ser atractivos de cara al cliente y no generar sobre endeudamiento en los individuos.

Con ello en mente, los cuatro elementos básicos para el desarrollo del modelo de negocios propuesto con este proyecto son:

- **Propuesta de valor para el cliente:** la propuesta de valor de cara al cliente es la de establecer un límite óptimo de endeudamiento que permita establecer una relación de largo plazo de ganar-ganar entre cliente y empresa, así como también generar precios acorde al nivel de riesgo deseado para cada uno de ellos, de acuerdo a sus características.
- **Fórmula de utilidades:** para incrementar las utilidades, lo que se pretende principalmente es reducir los costos por tres vías: primero, al tener un modelo de predicción de gasto en riesgo, el presupuesto se puede asignar de manera más eficiente;

⁶ (Johnson, Christensen, & Kageman, 2008).

⁷ El *spread* corresponde a la diferencia entre la tasa de compra de los activos financieros y la tasa de venta de los mismos. Representa una medida de ganancia de las colocaciones del Banco, antes de gastos operacionales y de riesgo.

segundo, dado que el perfil de riesgo mejora al no sobre endeudar a los clientes, también mejora el gasto en provisiones por riesgo de crédito; y tercero, los precios mínimos asegurarán que se compense el gasto en riesgo asociado a cada venta de créditos vía campaña.

- **Recursos clave:** los recursos clave involucrados en el proyecto son mínimos: si bien se cuenta con un *data ware house* corporativo (DWH), la información se encuentra desagregada y en formato poco legible. Para llevar a cabo la construcción de los modelos que se pretende implementar, se construyó un *data mart* con información específica de riesgo, la que se encuentra estructurada y almacenada históricamente.
- **Procesos clave:** el proceso clave es el proceso de venta. Sin embargo, no se abarcará el proceso completo, sino una parte de él, la que tiene relación con la administración de la relación con el cliente.

2.10 Planteamiento estratégico

De acuerdo al modelo Delta de Hax⁸, el proyecto busca potenciar a Bci estratégicamente en torno a la dimensión del mejor producto. Para ello, el posicionamiento que se desea alcanzar es el de liderazgo en costos, entregando servicios financieros casi personalizados (en términos de cupos) y con precios únicos de acuerdo al nivel de riesgo que se espera y desea alcanzar con cada cliente. Esto impacta directamente en la fórmula de utilidades descrita más atrás.

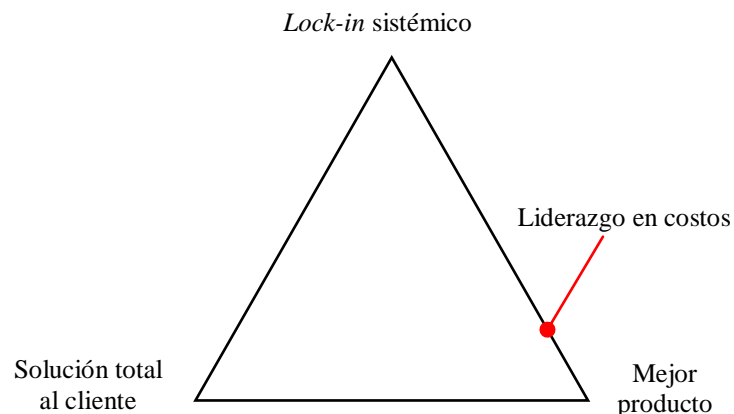


Imagen 9: Posicionamiento estratégico de acuerdo a modelo Delta

⁸ (Hax, 2010).

2.11 Impacto del Proyecto en la Estrategia Actual

La estrategia corporativa del Banco, tal como se vio más atrás, considera 5 ámbitos clave: colaboradores, innovación, procesos, cliente y financiero. El proyecto impactará en todos estos ámbitos clave de la siguiente manera:

- **Colaboradores:** existen participantes en el proceso de generación de campañas que destinan mucho tiempo en el cálculo de las nóminas. Al automatizar estas tareas, este tiempo se libera, dejando más espacio para otras tareas, como análisis.
- **Innovación:** Bci se define como un banco innovador. En este sentido, la propuesta del proyecto incide directamente en este ámbito al implementar metodologías de uso de analítica, técnicas que actualmente se utilizan someramente.
- **Procesos:** el proyecto rediseñará el proceso de generación de campañas de créditos de consumo, que está anidado dentro del proceso de venta de créditos de consumo vía campañas, con el compromiso de lograr implementar procesos *lean*.
- **Cliente:** el cliente se verá beneficiado al contar con un servicio de servicios financieros más adecuado a su realidad (no sobre endeudándolo) y cobrándole un precio adecuado. Adicionalmente, una posible mejor política de administración de riesgo podría eventualmente abrir las puertas de Bci a clientes que antes no podían ser atendidos de forma masiva.
- **Financiero:** el Banco se verá beneficiado de dos maneras: primero, por disminuir los costos operacionales, y segundo, por establecer mejores perfiles de riesgo de la cartera para el flujo. Adicionalmente, las pérdidas actuales por créditos mal colocados se disminuirá, generando una fuente adicional de utilidad.

2.12 Resumen del planteamiento del proyecto

De acuerdo a la metodología *Lean Change Method*⁹, podemos resumir gran parte de lo mencionado anteriormente según la metodología *canvas*. Esta metodología consiste en colocar en un lienzo los distintos componentes principales de un proyecto de cambio y trazar la estrategia del proyecto en torno al objetivo definido. Según esta metodología, estos componentes (ilustrados en la Imagen 10) son:

- **Urgencia:** el problema actual y por el que se gesta el proyecto, se divide en las siguientes ramas. Primero, los modelos actuales de riesgo predicen el gasto en riesgo de la cartera como un stock y se recalculan mensualmente. Aun cuando sirven para tomar decisiones respecto de la administración de la cartera, los modelos actuales son poco predictivos en el gasto real de una colocación nueva, tendiendo a sobre estimar el gasto en riesgo. Segundo, los precios no están alineados con el gasto en riesgo. En algunas operaciones

⁹ (Anderson, 2013).

se obtienen utilidades y en otras, pérdidas. No hay precios diferenciados por nivel de riesgo esperado. Tercero, el proceso de generación de campañas masivas es totalmente manual, tomando un tiempo humano excesivo en tareas que no aportan valor necesariamente y que son totalmente automatizables.



Imagen 10: Canvas del proyecto

- **Estado deseado:** ya identificado el problema actual, la situación deseada considera procesos de generación de campañas automáticos, precios diferenciados por riesgo y límites de endeudamiento óptimos de acuerdo al riesgo esperado.
- **Acciones:** las tareas que se deben realizar para lograr el estado deseado tienen que ver con el rediseño del proceso utilizando la metodología del MBE, y el desarrollo de los modelos utilizando la metodología KDD. Para la aplicación del proyecto se ejecutará un piloto, tomando un porcentaje aleatorio de la cartera y aplicando el modelo (grupo de tratamiento o estrategia *challenger*) y se contrastará el desempeño con respecto del resto de los clientes donde no se aplicó el modelo (grupo de control o estrategia *champion*).
- **Visión:** la visión del proyecto es la del PTR, que consiste en empapar la cultura de riesgo en toda la organización, siendo consistente con los criterios de experiencia de Bci.
- **Comunicación:** el proyecto cuenta con un plan de gestión del cambio (que veremos en detalle más adelante). Dos de los aspectos más relevantes del plan de gestión del cambio

tienen que ver con establecer una serie de reuniones donde se revisarán los avances del proyecto, y la medición periódica (semanal) de los distintos KPIs establecidos para medir el éxito.

- ***Criterios de éxito:*** como medida de desempeño, habrá 3 grupos de KPIs: relacionados al proceso (tiempo de procesamiento, liberación de recursos, cumplimiento de compromisos), relacionados a los modelos (estabilidad de los modelos, medidas de error, medidas de gasto en riesgo aproximadas por mora) y relacionados al proyecto global (tasa de créditos cursados sobre el total de clientes en campaña, para el grupo de tratamiento y el grupo de control).
- ***Receptores del cambio:*** todos los involucrados en el proceso de venta se verán afectados en menor o mayor medida. Se ha identificado dos grupos de participantes, de acuerdo al nivel de intervención: relacionados con la generación de las campañas (principalmente unidades relacionadas con la Gerencia de Riesgo) y relacionados con las actividades anexas (principalmente Gerencias Comerciales, de Marketing y de Canales).
- ***Fórmula de utilidades:*** los costos asociados al proyecto tienen que ver con el uso del tiempo (en horas-hombre, HH) de quienes rediseñarán en proyecto, programarán la solución tecnológica y desarrollarán los modelos. Existe un costo adicional, aunque hundido, ya que se trata de recursos existentes en la organización, relacionado con el costo de las licencias de software, uso de servidores y otros. En relación a los ingresos del proyecto, podemos mencionar que se obtendrá ganancia producto de precios mejor asignados (asegurando *spread* mínimos deseados), tiempo de los ejecutores liberado para otras tareas (como análisis) y procesos *lean*. En la sección siguiente se detallará con mayor detenimiento la fórmula de utilidades.

2.13 Justificación económica del proyecto

2.13.1 Antecedentes para la evaluación

El proyecto será evaluado utilizando la metodología de evaluación tradicional de proyectos, con los indicadores clásicos de esta metodología (flujo de caja, valor actual neto –VAN– y tasa interna de retorno –TIR).

Para la evaluación económica del proyecto se deberá considerar la inversión inicial, el flujo de caja y la tasa de descuento inter temporal. La inversión inicial la constituye:

- Infraestructura tecnológica destinada a realizar tareas de servidor de datos y servidor de aplicaciones (a nivel de prototipo, estos equipos coinciden con un equipo disponible en la unidad, que no se está utilizando más que para funciones de respaldo de información;

coincidentalmente, estos equipos cuentan con las licencias de software compatibles con el prototipo).

- Tiempo destinado a la planificación, diseño y ejecución del proyecto, medido en horas hombres (y traducible a efectivo prorrateando parte de la renta de las personas involucradas).
- Tiempo destinado al desarrollo de los modelos y su aplicación, medido en horas hombres.

Para una correcta contabilización de los flujos, se deberá considerar los siguientes costos:

- Costo de mantenimiento del sistema tecnológico de apoyo.
- Costo de ejecución de los modelos de venta.

Y como potenciales ingresos, se contabiliza lo siguiente:

- Ganancia por mejor ajuste de los precios a cobrar.
- Ganancia por potencial mayor venta.
- Ganancia por liberación de provisiones constituidas para el flujo de créditos vendidos a través de campañas.

Adicionalmente habrá otro tipo de efectos producto del proyecto, que será más difícil de contabilizar, y que en una primera instancia sólo serán enunciados:

- Producto de una mejor asignación de los límites de crédito para clientes, las políticas de riesgo podrían ser modificadas, impactando en el total de la cartera y principalmente en el costo utilizado en evaluar caso a caso en los comités de créditos.
- Producto de la liberación de tiempo dedicado a tareas de proceso, el ejecutor del proceso puede generar nuevas oportunidades para el banco, que pueden generar costos e ingresos de difícil cuantificación.
- Ganancias por reinversión del ahorro en provisiones.
- Posible rediseño de la estructura organizacional.
- Otros efectos no descubiertos aún.

Respecto del horizonte de evaluación y la tasa de descuento inter temporal a utilizar, se considerará un plazo de evaluación de 2 años, con periodicidad de medición de los flujos semestral, que corresponde al tiempo en que las políticas y procesos de riesgo se actualizan. La tasa de descuento inter temporal será la que utiliza Bci en sus estudios de proyectos.

2.13.2 Construcción del Flujo de Caja

Ante lo mencionado anteriormente, expresaremos brevemente los costos, ingresos e inversión asociados al proyecto. Para todos los efectos, los costos e ingresos que acá se presentan son incrementales respecto de la situación sin proyecto (donde se aplicaron ya algunas mejoras en el proceso, pero no se implementaron los modelos de riesgo).

2.13.2.1 *Inversión inicial*

La inversión inicial considerará un plazo máximo de un semestre. Se requerirá para la puesta en marcha del proyecto un equipo (a utilizar como servidor de aplicaciones y de datos), así como también horas hombre (HH) de distintos actores. Estas horas hombre ya están valorizadas a momento cero, es decir, no se amortizan a lo largo de la vida del proyecto. La Tabla 2 muestra la inversión inicial considerada.

Las HH de los actores considerados como inversión se justifica por el uso de:

- Ejecutivo de Estrategias (1): es quien lidera el proyecto y desarrolla los modelos y el prototipo computacional. Dedicará el 40% de su tiempo durante 6 meses.
- Sub Gerente de Estrategias: es quien hace las funciones de sponsor del proyecto, y sus funciones serán las de validar el proyecto y controlar que los hitos se ajusten a lo esperado. Dedicará el 10% de su tiempo durante 3 meses.
- Ejecutivo de Estrategias (2): será necesario contar con la ayuda de un segundo ejecutivo de estrategias, quien estará a cargo de ayudar en las tareas de automatización del proceso de generación de campañas.

En lo referente al hardware, se considera una tasa de uso del 10% de un equipo con capacidad ociosa sólo por efectos de depreciación lineal, calculado a una vida útil de 5 años (10 semestres) sin valor residual. No se considera el costo de las licencias, ya que no son exclusivas del proyecto (fueron adquiridas previamente para tareas propias del negocio).

Ítem	Meses	Uso	Costo unitario	Inversión total
Equipo	3	10%	M\$ 750	M\$ 225
HH Ejecutivo de Estrategias (1)	6	40%	M\$ 1,600	M\$ 3,840
HH Sub Gerente Estrategias	6	10%	M\$ 4,000	M\$ 2,400
HH Ejecutivo de Estrategias (2)	1	40%	M\$ 1,500	M\$ 600
Total inversión				M\$ 7,065

Tabla 2: Inversión inicial considerada en el proyecto

2.13.2.2 *Costos proyectados*

Los egresos del proyecto se dividieron en costos directos y gastos de administración. Dentro de los costos directos, se considera las horas hombre de dos ejecutivos de estrategias: el primero (1) para labores de mantención de los modelos y seguimiento; el segundo (2) para tareas de mantención del proceso y el sistema. El uso del primero será decreciente en el tiempo, a excepción del último semestre, donde se realizará una mantención exhaustiva, a modo de determinar la vigencia del proyecto (y en caso contrario, levantar un nuevo proyecto para la reformulación de los modelos y el sistema). El uso del segundo será decreciente en el tiempo, producto de que se espera alcanzar madurez en el uso del proceso y el sistema. Otro costo relacionado clasificado como directo tiene que ver con la mantención de los equipos, que para efectos prácticos se considerará lineal y equivale a mantención de software y hardware, equivalente en monto al valor de la depreciación.

Por otra parte, dentro de los gastos de administración, se considera un plan de comunicación durante el primer semestre de vida del proyecto, así como también un plan de capacitación permanente (pero decreciente en uso en el tiempo).

Ambos egresos se muestran en la Tabla 3.

Ítem	Costo unitario	Semestre 1		Semestre 2		Semestre 3		Semestre 4	
		Uso	Inversión total	Uso	Inversión total	Uso	Inversión total	Uso	Inversión total
HH Ejec. Estrategias (1)	M\$ 1,600	20%	M\$ 320	10%	M\$ 160	10%	M\$ 160	50%	M\$ 800
HH Ejec. Estrategias (2)	M\$ 1,500	30%	M\$ 450	25%	M\$ 375	20%	M\$ 300	15%	M\$ 225
Mantenición de equipos			M\$ 225		M\$ 225		M\$ 225		M\$ 225
Total costos directos			M\$ 995		M\$ 760		M\$ 685		M\$1,250
Plan de comunicación	M\$ 1,800	20%	M\$ 360		M\$ -		M\$ -		M\$ -
Capacitación	M\$ 1,500	20%	M\$ 300	5%	M\$ 75	5%	M\$ 75	1%	M\$ 15
Total gastos de administración			M\$ 660		M\$ 75		M\$ 75		M\$ 15
Total costos proyectados			M\$1,655		M\$ 835		M\$ 760		M\$1,265

Tabla 3: Costos esperados con el proyecto

2.13.2.3 Ingresos proyectados

Los principales ingresos se simularon considerando tres escenarios: uno pesimista, uno optimista y uno más probable. Las fuentes de ingreso identificadas en este proyecto son:

- (i) Mejor asignación de precios: al cobrar precios ajustados por riesgo se elimina la pérdida por precios mal cobrados y se asegura un *spread* mínimo.
- (ii) Ajuste de venta por nichos: al tener modelos predictivos del gasto en riesgo, se espera colocar en nichos donde antes no se ofrecía y se espera reducir la venta en nichos más riesgosos. Se espera que los nuevos modelos predigan una venta mejor ajustada, redistribuyendo la venta de acuerdo al siguiente cuadro:

	Incremento venta (MM\$)
E1: Escenario conservador (10% de la cartera)	7,860
E2: Escenario más probable (90% de la cartera)	70,838
E3: Escenario optimista (100% de la cartera)	78,698

Tabla 4: Ingresos esperados por mejor asignación de leads en campaña

Fuente: elaboración propia con datos de Bci.

- (iii) Liberación de provisiones: ahorro por liberación de provisiones constituidas para el flujo de créditos vendidos a través de campañas. Este es un efecto secundario

no atribuible directamente al proyecto. Para ser más conservadores, este potencial ingreso no se considerará en la evaluación económica.

- (iv) Ahorro de recursos humanos: liberación de tiempo humano en ejecución del proceso de generación de campañas. Este ingreso potencial es más difícil de medir, y al igual que en el caso anterior, se dejará fuera de la evaluación.

Estos potenciales ingresos se pueden revisar en la Tabla 5.

	Ingreso potencial (MM\$)
Ganancia por ajuste de precios	2,685
Ganancia por ajuste de venta - E1	7,860
Ganancia por ajuste de venta - E2	70,838
Ganancia por ajuste de venta - E3	78,698

Tabla 5: Ingresos esperados con el proyecto

2.13.2.4 Flujo de caja

Para la evaluación, se considera una tasa de descuento del 5.88% por semestre¹⁰ y un horizonte de evaluación de 4 semestres (2 años).

Luego, para el escenario conservador, el flujo de caja es el siguiente:

	Semestre 0	Semestre 1	Semestre 2	Semestre 3	Semestre 4
Inversión inicial	-7,065				
Ingresos operacionales		3,286,290	3,286,290	3,286,290	3,286,290
Costos operacionales		-995	-760	-685	-1,250
Gastos de administración		-660	-75	-75	-15
Resultado operacional		3,284,635	3,285,455	3,285,530	3,285,025
Ingresos no operacionales		0	0	0	0
Costos no operacionales		0	0	0	0
Resultado no operacional		0	0	0	0
Utilidad antes de impuesto	-7,065	3,284,635	3,285,455	3,285,530	3,285,025
Impuesto (21%)		-689,773	-689,946	-689,961	-689,855
Utilidad después de impuesto		2,594,862	2,595,509	2,595,569	2,595,170
Flujo de caja libre	-7,065	2,594,862	2,595,509	2,595,569	2,595,170

Tabla 6: Flujo de caja, escenario conservador (todos los valores en M\$)

Este flujo de caja implica que el Valor Actual Neto del proyecto es de MM\$8,510, con una Tasa Interna de Retorno del 36,728.43%.

¹⁰ Entregado por Bci estudios.

Al revisar el escenario más probable, el flujo de caja es el siguiente:

	Semestre 0	Semestre 1	Semestre 2	Semestre 3	Semestre 4
Inversión inicial	-7,065				
Ingresos operacionales		8,104,107	8,104,107	8,104,107	8,104,107
Costos operacionales		-995	-760	-685	-1,250
Gastos de administración		-660	-75	-75	-15
Resultado operacional		8,102,452	8,103,272	8,103,347	8,102,842
Ingresos no operacionales		0	0	0	0
Costos no operacionales		0	0	0	0
Resultado no operacional		0	0	0	0
Utilidad antes de impuesto	-7,065	8,102,452	8,103,272	8,103,347	8,102,842
Impuesto (21%)		-1,701,515	-1,701,687	-1,701,703	-1,701,597
Utilidad después de impuesto		6,400,937	6,401,585	6,401,644	6,401,245
Flujo de caja libre	-7,065	6,400,937	6,401,585	6,401,644	6,401,245

Tabla 7: Flujo de caja, escenario más probable

Con este flujo de caja, el VAN del proyecto es de MM\$21,001, con una TIR del 90,600.68%.

Finalmente, en el más optimista de los escenarios, el flujo de caja es el siguiente:

	Semestre 0	Semestre 1	Semestre 2	Semestre 3	Semestre 4
Inversión inicial	-7,065				
Ingresos operacionales		8,705,397	8,705,397	8,705,397	8,705,397
Costos operacionales		-995	-760	-685	-1,250
Gastos de administración		-660	-75	-75	-15
Resultado operacional		8,703,742	8,704,562	8,704,637	8,704,132
Ingresos no operacionales		0	0	0	0
Costos no operacionales		0	0	0	0
Resultado no operacional		0	0	0	0
Utilidad antes de impuesto	-7,065	8,703,742	8,704,562	8,704,637	8,704,132
Impuesto (21%)		-1,827,786	-1,827,958	-1,827,974	-1,827,868
Utilidad después de impuesto		6,875,956	6,876,604	6,876,663	6,876,264
Flujo de caja libre	-7,065	6,875,956	6,876,604	6,876,663	6,876,264

Tabla 8: Flujo de caja, escenario optimista

Este escenario presenta un VAN de MM\$22,559 con una TIR del 97,324.23%.

En cualquiera de estos escenarios, los indicadores de evaluación de proyectos indican que es rentable, dada la baja inversión y los altos ingresos obtenidos en un muy corto plazo (ingresos inmediatos).

2.14 Gestión del cambio del proyecto

El proyecto de tesis tiene asociado un proyecto de cambio que pretende darle sustento desde el punto de vista humano en la organización en la que está inmerso. Para darle sentido al proyecto de cambio, repasaremos brevemente algunas distinciones más relevantes que estarán presentes a lo largo del proyecto de tesis.

2.14.1.1 *Sentido de urgencia*

El proyecto de tesis tiene que tener un sentido de urgencia que le dé importancia. El sentido de urgencia que se le da al presente proyecto de tesis es el de predecir el nivel de gasto en riesgo, toda vez que se quiere liberar gasto en riesgo. Esto se logrará al asignar precios correctos y límites de endeudamiento adecuados. Adicionalmente, el proceso actual se ejecuta de forma manual, utilizando mucho tiempo humano en tareas que no agregan valor. Es necesario disminuir este tiempo.

2.14.1.2 *Visión del proyecto*

El proyecto actual está enmarcado en el Plan de Transformación de Riesgo, por lo que la visión del proyecto se alinea en torno a la visión del PTR: “Buscamos desarrollar una sólida cultura de riesgo, a nivel organizacional, para hacer realidad los sueños de nuestros clientes, adoptando permanentemente las mejores prácticas de gestión de riesgo, a través de líderes y equipos comprometidos, que nos permitan ser reconocidos a nivel nacional y regional como el banco con la mejor relación riesgo/retorno”.

2.14.1.3 *Identificación de los stakeholders y su gestión*

Los principales *stakeholders* relacionados al proyecto son:

- Gerente de Riesgo y jefaturas dependientes de él: el interés principal radica en que se podrá medir e imputar de mejor manera el gasto en riesgo de la venta de créditos de consumo, establecer potenciales políticas de crédito multidimensionales y posiblemente (en el caso de un *roll out* del piloto) procesos de aprobación más expeditos. Su gestión será la de aprobadores y evaluadores de la nueva metodología y los pilotos.
- Sub Gerente de Pricing: su interés radica en el establecimiento de precios adecuados al nivel de riesgo de cada cliente, lo que implica que su gestión tiene que ver con la capacidad de volcar los modelos en las políticas de fijación de precios y facultamientos de excepciones.
- Gerencias Comerciales: al impactar directamente en las oportunidades de negocio que la Gerencia de Riesgo disponibiliza a través de campañas masivas, el interés de los distintos Gerentes Comerciales es total. Su gestión tiene relación con fomentar la venta de clientes en el piloto.

2.14.1.4 Cambio y conservación

Antes de poner en marcha el proyecto hay que tener presente qué es lo que cambiará y lo que se conservará luego de su implementación. Lo que cambiará es la forma de ejecutar los procesos al interior de la organización, pero principalmente una forma distinta de evaluar a clientes. Las metodologías actuales ven dos o tres dimensiones de los clientes en la evaluación de su potencial nivel de endeudamiento. La nueva metodología propuesta en esta tesis ya no ve a un cliente como un ente de dos o tres dimensiones, sino que se incorporan otros atributos, introduciendo riqueza en la forma de evaluar a los clientes y mayor completitud.

2.14.1.5 Plan de comunicación

Tal como se adelantó, se llevará un plan comunicacional intensivo desde el inicio del proyecto. Se consensuará el rediseño de los procesos de generación de campañas, se informará a los otros participantes del proceso de venta la importancia de estos cambios, entre otras. Para ello, se utilizarán los canales de comunicación formales que el Banco dispone.

En el plan de comunicación también se contempla llevar a cabo reuniones periódicas con los principales *stakeholders* para informar no tan solo el avance del proyecto, sino también nuevas iniciativas que se puedan implementar en el transcurso del desarrollo del proyecto de tesis.

2.14.1.6 Potencial para desarrollo futuro

Al implementar un piloto para probar este proyecto como una estrategia *challenger*, queda abierta la posibilidad de que, una vez terminada la etapa de piloto, el *challenger* pase a modificar la política de créditos. Esto daría pie a un posible futuro trabajo para hacer el *roll out* del piloto al 100% de la cartera *Retail*. Esto podría significar oportunidades de desarrollo laboral al interior de la organización, tanto para el autor como para otros participantes.

3 ARQUITECTURA DE PROCESOS

3.1 Patrón de negocios

De acuerdo a la metodología propuesta por Barros¹¹, la *Arquitectura de Macro Procesos* se enmarca dentro de un patrón más amplio desde el cual se puede modelar el funcionamiento de los procesos internos de la organización.

Barros en su metodología propone una serie de *patrones de negocios* que sirven de marco de referencia para el desarrollo de cualquier rediseño de procesos o de diseño de procesos nuevos. Por lo tanto, este proyecto se enmarca dentro del patrón de negocios 1 (BP1). Este patrón de negocios explica cómo se justifica este proyecto: partiendo de la cadena de valor descrita anteriormente. Se revisa la relación con el cliente, revisión que se monitorea constantemente. De esta revisión surgen requerimientos de negocios para modelar la cadena de valor de una manera más eficiente. Producto de esta revisión, se diseñarán procesos nuevos para la creación y aplicación de modelos de atención para la venta, y se rediseñarán los procesos de venta existentes, con el fin de interiorizar los lineamientos de venta en las áreas que ejecutan esta acción. Posteriormente, junto con la generación de ofertas, se entregan ofertas de campaña mejor diseñadas y sustentadas en procesos más limpios. Todo el proceso se apoya en las distintas áreas de apoyo del Banco. El BP1 asociado a este proyecto se puede revisar en la Imagen 11.

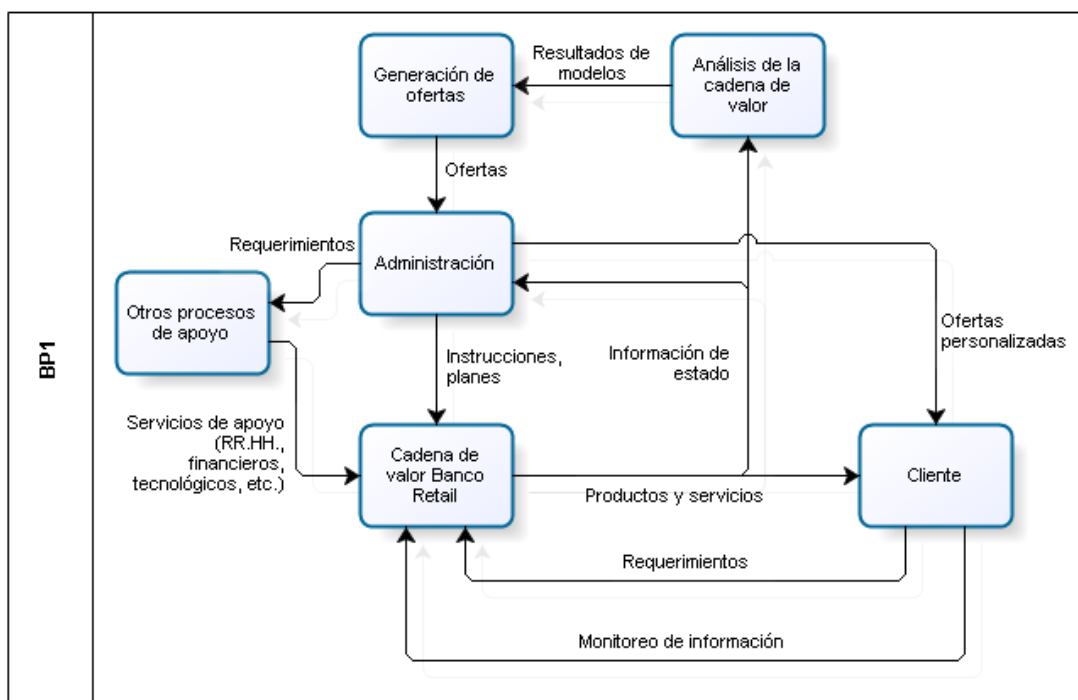


Imagen 11 Patrón de arquitectura de procesos BP1

¹¹ (Barros, 2012)

3.2 Variables de diseño

El patrón de negocios que sustenta el proyecto tiene como fundamentos una serie de variables de diseño que le dan vida. Estas variables de diseño se agrupan en distintos aspectos: estructura de la empresa y mercados, coordinación, prácticas de trabajo, integración de procesos y mantención consolidada de estados.

En lo referente a la estructura de la empresa y mercados, las variables de diseño que permiten la implementación de este proyecto son las descritas en la Tabla 9:

	Situación actual	Situación propuesta
Estructura interna	Mucha participación manual	Automatización del proceso implica liberar tiempo de sus participantes
Toma de decisiones	Descentralizada, descoordinación entre las políticas de crédito y su aplicación en campañas masivas	Gerencia de Riesgo, unidad de Políticas de Crédito

Tabla 9: Variables de diseño: estructura de la empresa y mercados

Respecto a la coordinación, las siguientes son las diferencias entre la situación actual y el escenario propuesto con este proyecto:

	Situación actual	Situación propuesta
Reglas	Reglas formales, pero desalineadas con la política de créditos	Centralización de las reglas del proceso en una sola unidad (controles)
Jerarquía	Aprobación de Gerente de Riesgo para aplicación de pilotos, autorización del directorio para políticas nuevas	
Partición y colaboración	Existe mucha intervención entre participantes (riesgo operacional elevado)	Se redistribuyen funciones entre participantes y se automatizan tareas críticas (mitigando riesgo operacional)

Tabla 10: Variables de diseño: coordinación

Las prácticas de trabajo se pueden resumir en la Tabla 11:

	Situación actual	Situación propuesta
Lógica automatizada o semi automatizada	No automatizada	100% automatizada
Lógica de apoyo a actividades tácitas	No declarada	Bitácora de modificaciones
Procedimientos de comunicación e integración	Normativas incongruentes en una sola plataforma	Normativa unificada en una sola plataforma

Tabla 11: Variables de diseño: prácticas de trabajo

La integración de procesos conexos se modificará de acuerdo a lo planteado en la Tabla 12:

	Situación actual	Situación propuesta
Existencia de proceso	Proceso existente, pero informalmente	Un solo proceso formal
Todos o la mayor parte de procesos en un macro proceso	Sí, en macro 1	

Tabla 12: Variables de diseño: integración de procesos conexos

Finalmente, la mantención consolidada de estados presentará las siguientes consideraciones:

	Situación actual	Situación propuesta
Datos propios	Parcialmente	Repositorios independientes (<i>data mart</i>) y autónomos
Integración con datos de otros sistemas	Integración (manual) con diversos sistemas	No es necesaria

Tabla 13: Variables de diseño: mantención consolidada de estados

3.3 Arquitectura de macro procesos

El proyecto que aquí se presenta se enmarca dentro del modelo de macro procesos desarrollado por Barros en el *Macro proceso 1*, tal y como se puede apreciar en la Imagen 12. Por simplicidad, el diagrama de proceso se ha resumido, y no se muestran otras cadenas de valor que comparten servicios entre sí, como la Cadena de Valor de venta de otros productos de créditos a clientes o la cadena de valor de venta de servicios financieros a nuevos clientes.

Al seguir desarrollando este patrón de macro procesos, podemos ver que el patrón en el que mejor encaja el proyecto tiene que ver con Administración de la relación con el cliente, tal y como se aprecia en la Imagen 14. Dentro del proceso de Administración de la relación con el cliente, el proyecto se enfocará directamente en Marketing y análisis de clientes. Este proceso puede ser visto en la Imagen 15, donde nos enfocaremos en el proceso Gestión de modelos de venta basados en riesgo (Imagen 16). El desarrollo de los modelos de originación se llevará a cabo en esta actividad.

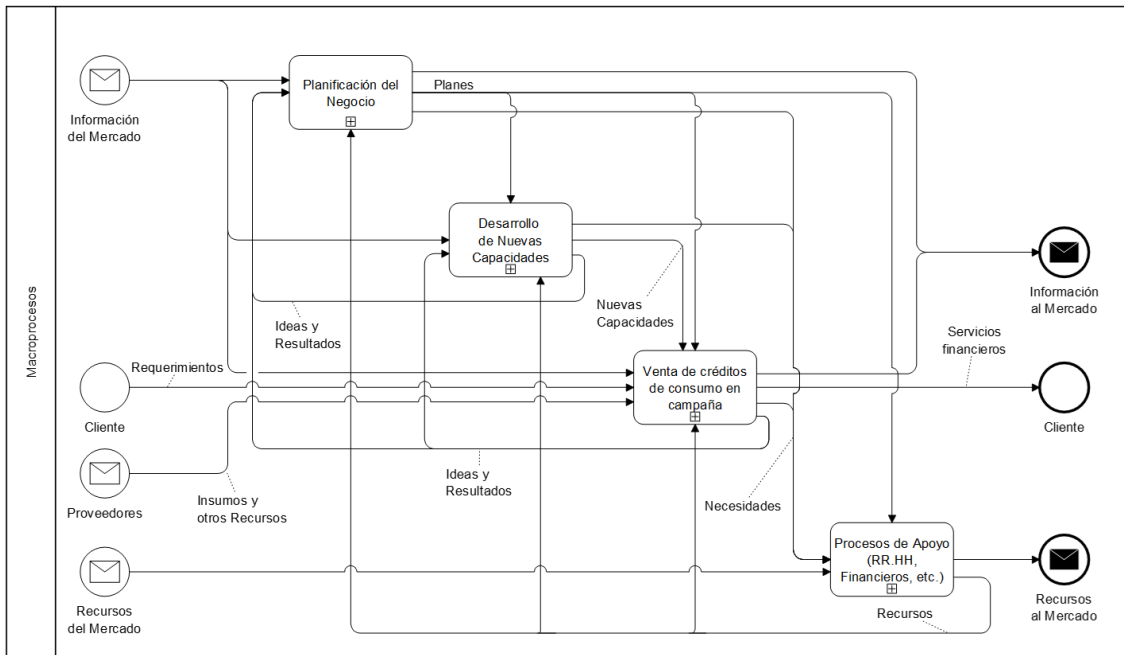


Imagen 12: Arquitectura de macro procesos del proyecto

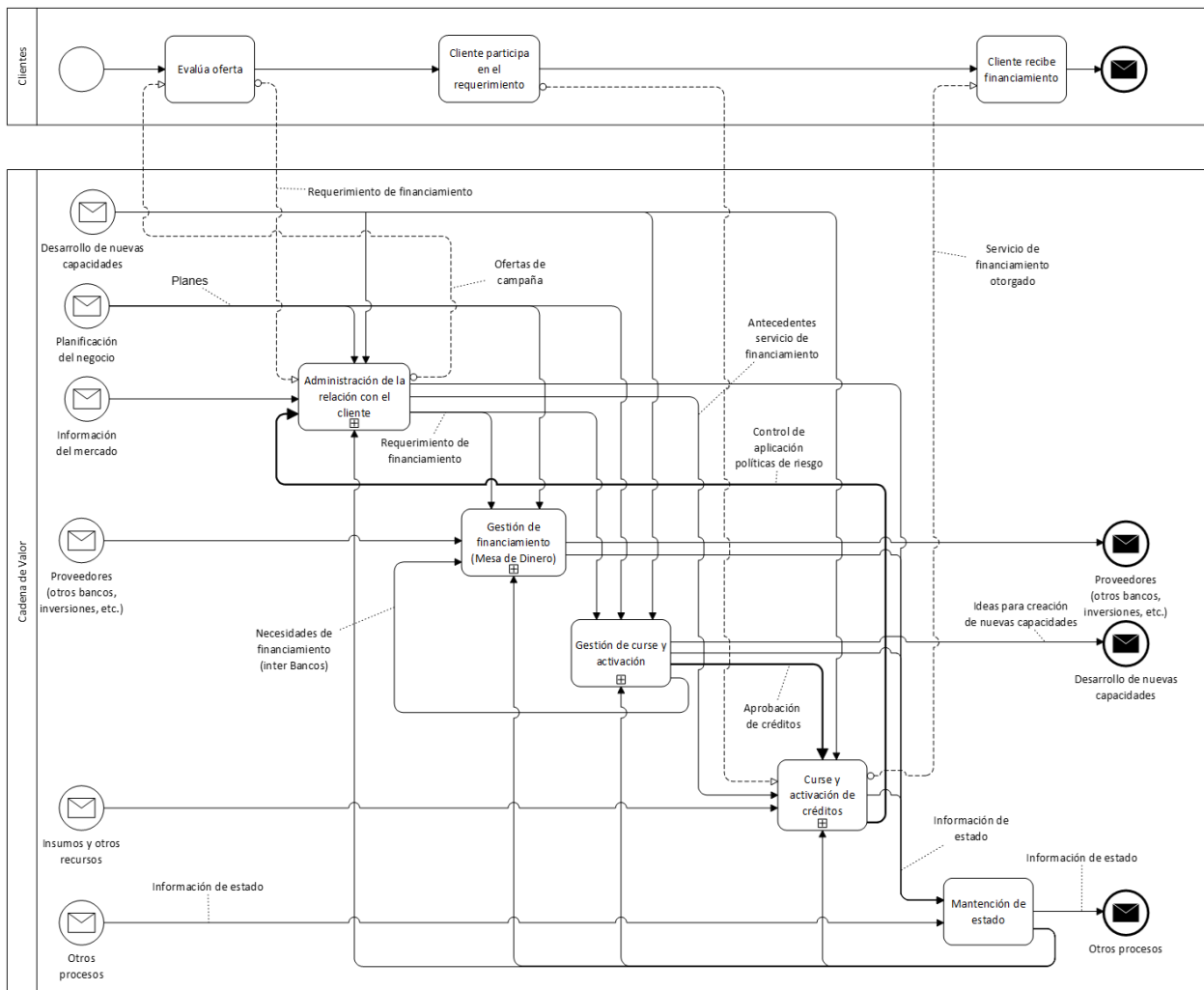


Imagen 13: Detalle de procesos, macro 1

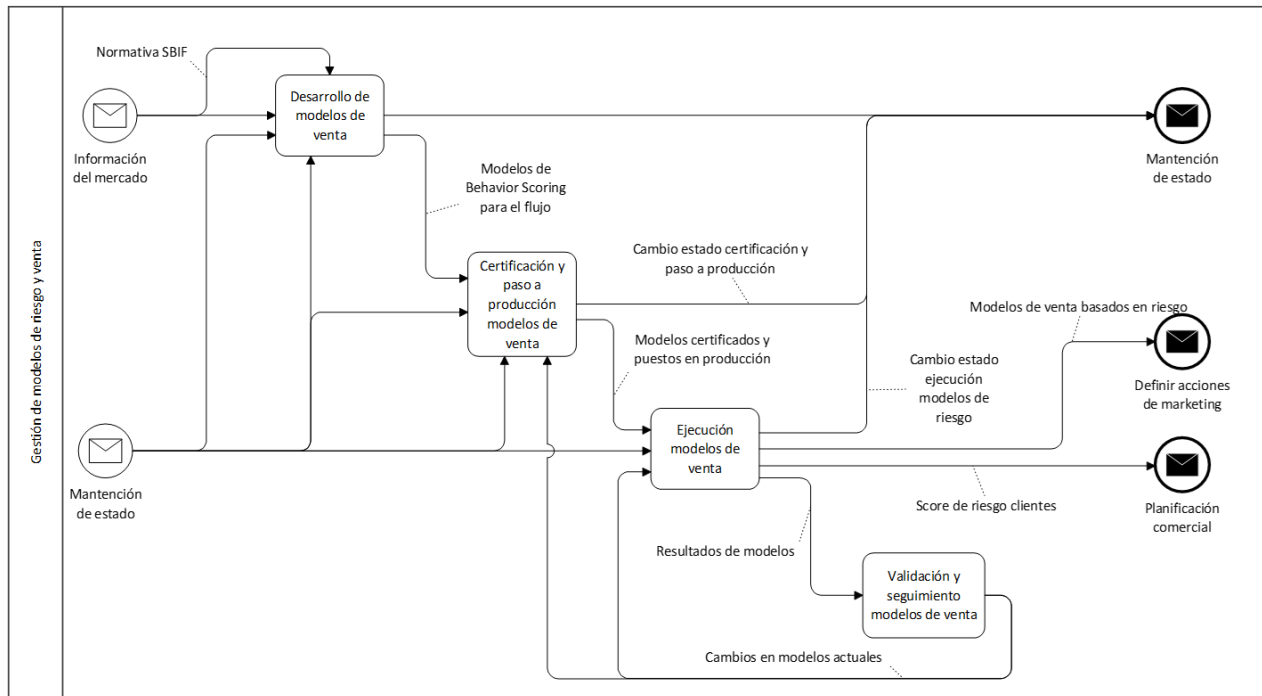


Imagen 16: Gestión de modelos de venta basados en riesgo

3.4 Diagramas de pistas

El proyecto de tesis se centra en las actividades de *Desarrollo de modelos de venta* y *Ejecución de modelos de venta*. Para la primera actividad, la metodología a emplear será *Knowledge Discovery in Databases (KDD)*, que resume una serie de pasos a seguir para buscar patrones que permitan encontrar conocimiento oculto en los datos, y que apoye la toma de decisiones. Para el segundo, se aplicarán técnicas del MBE y *lean change*, que serán detalladas a continuación.

3.4.1 Desarrollo de modelos de venta

La búsqueda de patrones de acuerdo a esta metodología debe comenzar con una recopilación de datos, para luego consolidarlos y prepararlos. El tercer paso es la búsqueda de patrones (que deben ser no triviales, nuevos, potencialmente útiles, escalables y replicables en el tiempo). El cuarto y último paso es el de consolidar e interpretar el conocimiento que se adquiera a través de estos patrones descubiertos, para así permitir la toma de decisiones al interior del negocio. Cada uno de estos pasos no es necesariamente secuencial, se puede ir iterando paulatinamente, con el fin de robustecer y mejorar la calidad de los patrones descubiertos.

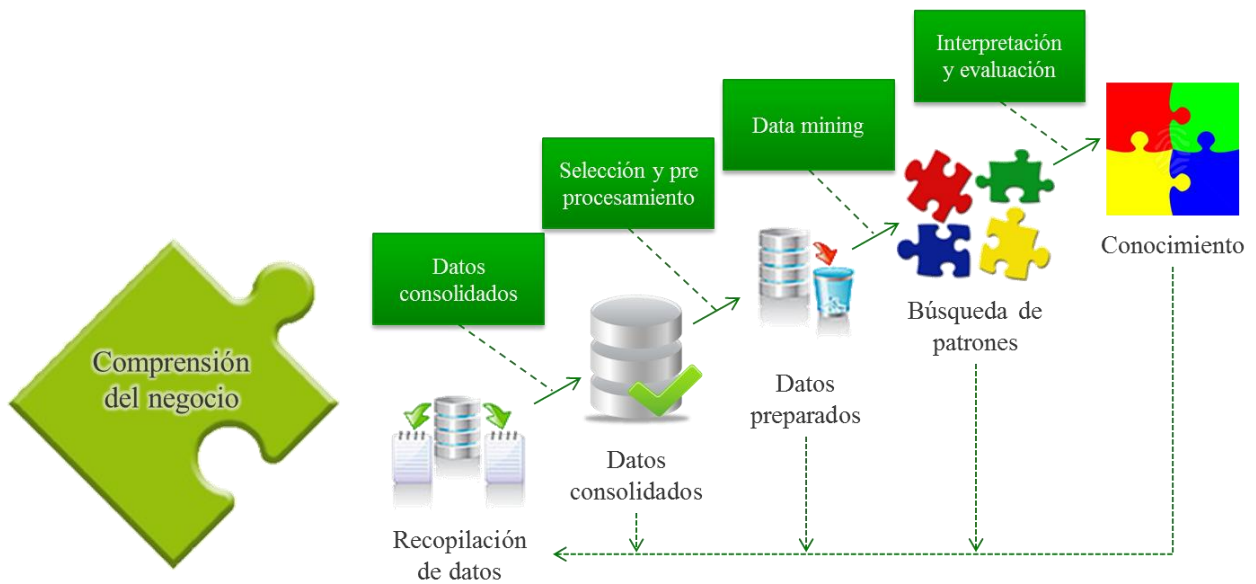


Imagen 17: Esquema de la metodología KDD seguida en el desarrollo de modelos de BI

Esta actividad del proceso comienza solicitando la extracción de datos para ejecutar el pre procesamiento. Luego, se ejecuta el resto de los pasos del modelo KDD para ambos escenarios. Al final del procesamiento de los modelos se evalúa el desempeño de los mismos. Si el desempeño, de acuerdo a criterios que se definirán más adelante, está conforme, los modelos se implantarán luego en el proceso Ejecución de modelos de venta. En caso que el desempeño no se ajuste a los niveles esperados por el analista de estrategias, decidirá si los modelos siguen siendo válidos y sólo se deben calibrar o si se deben reconstruir. La lógica de decisión se basará en la historia de ejecuciones y los indicadores de desempeño almacenados. En el caso que se decida que es suficiente una re calibración, se deberá repetir el proceso KDD, hasta lograr indicadores de desempeño dentro de los criterios que se consideran como aceptables. En caso contrario, los modelos deberán ser reevaluados por completo, cambiando los modelos ya existentes por unos nuevos. El resumen de este flujo puede ser revisado en la Imagen 18.

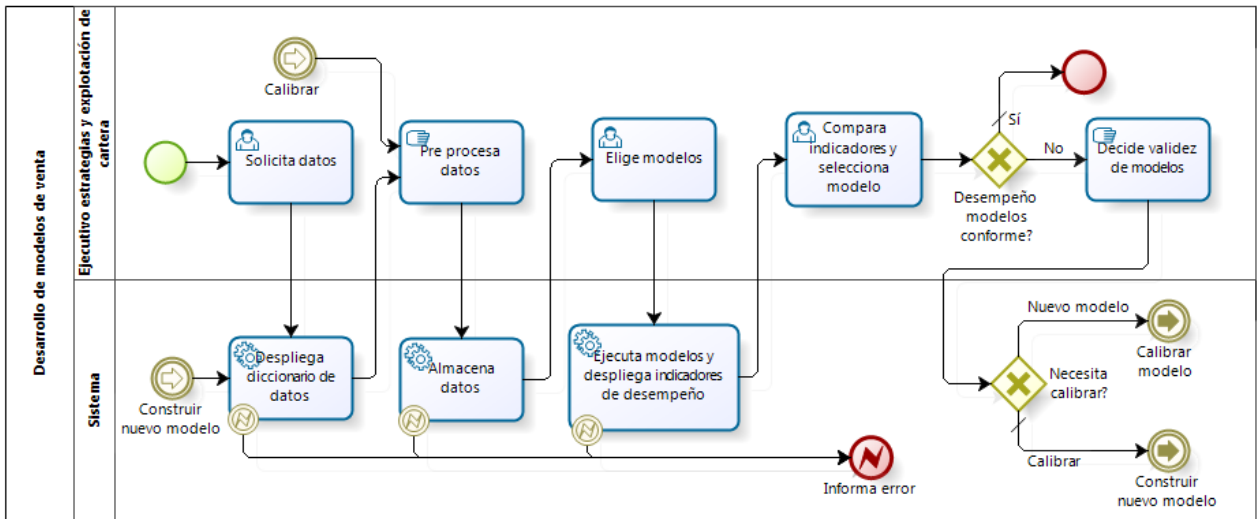








Imagen 18: Diagrama de pistas para Desarrollo de modelos de venta

El proceso se simuló en *Bizagi Process Modeler* v 2.6.0.4, para lo cual se estableció los siguientes parámetros para las actividades del proceso:

N°	Componente	Tipo de tarea	Duración/frecuencia de la tarea
1	 Solicita datos	De usuario	Duración constante: 5 minutos.
2	 Despliega diccionario de datos	De servicio	Duración constante: estándar 5 segundos. Tiene un evento de error adjunto, el cual ocurre según una distribución gamma con parámetros $k = 1$ y $\theta = 5$.
3	 Pre procesa datos	Manual	Duración distribuida uniformemente entre 180 y 300 minutos.
4	 Almacena datos	De servicio	Duración constante: 60 segundos. Tiene un evento de error adjunto, el cual ocurre según una distribución gamma con parámetros $k = 1$ y $\theta = 5$.
5	 Elige modelos	De usuario	Duración distribuida uniformemente entre 10 y 60 minutos.
6	 Ejecuta modelos y despliega indicadores de desempeño	De servicio	Duración distribuida uniformemente entre 30 a 45 minutos. Tiene un evento de error adjunto, el cual ocurre según una distribución gamma con parámetros $k = 1$ y $\theta = 5$.



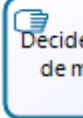

7	 Compara indicadores y selecciona modelo	De usuario	Duración distribuida uniformemente entre 60 y 90 minutos.
8	 Desempeño modelos conforme?	Puerta exclusiva	Frecuencia distribuida como: 80% de los casos desempeño conforme, 20% de los casos desempeño disconforme.
9	 Decide validez de modelos	Manual	Duración distribuida uniformemente entre 15 y 60 minutos.
10	 Necesita calibrar?	Puerta exclusiva	Frecuencia distribuida como: 95% de los casos sólo calibrar, 5% de los casos crear un nuevo modelo.

Tabla 14: Parámetros configurados para la simulación de Desarrollo y seguimiento de modelos de venta

Los resultados de la simulación del proceso se muestran en la Tabla 15. En ella se puede apreciar que en promedio, el proceso tardaría, de acuerdo a los parámetros configurados de acuerdo a lo expuesto anteriormente, 7 horas y 15 minutos aproximadamente.

Nombre	Tipo	Instancias completadas	Instancias iniciadas	Tiempo mínimo (m)	Tiempo máximo (m)	Tiempo prom. (m)	Tiempo total (m)
Desarrollo de modelos de venta	Proceso	1	2	434.34	434.34	434.34	434.34
NoneStart	Evento de inicio	2					
Solicita datos	Tarea	1	2	5	5	5	5
Despliega diccionario de datos	Tarea	1	1	0.08	0.08	0.08	0.08
Pre procesa datos	Tarea	1	1	289.23	289.23	289.23	289.23
Ejecuta modelos y despliega indicadores de desempeño	Tarea	1	1	38.86	38.86	38.86	38.86
Compara indicadores y selecciona modelo	Tarea	1	1	74.31	74.31	74.31	74.31
Elige modelos	Tarea	1	1	25.86	25.86	25.86	25.86
Desempeño modelos conforme?	Compuerta	1	1				

NoneEnd	Evento de Fin	1					
Necesita calibrar?	Compuerta	0	0				
Decide validez de modelos	Tarea	0	0	0	0	0	0
Construir nuevo modelo	Evento intermedio	0	0				
Calibrar	Evento intermedio	0	0				
Construir nuevo modelo	Evento intermedio	0	0				
Calibrar modelo	Evento intermedio	0	0				
Informa error	Evento de Fin	0					
Almacena datos	Tarea	1	1	1	1	1	1

Tabla 15: Resultados de la simulación de Desarrollo de modelos de venta

3.4.2 Ejecución de modelos de venta

El proceso, tal como se ilustra en la Imagen 19, comienza con la invocación de la ejecución de modelos de venta por parte del ejecutivo de estrategias y explotación de cartera. El sistema toma el requerimiento y ejecuta los modelos de venta (que fueron seleccionados en el proceso *Desarrollo de modelos de venta*) y los ejecuta para el universo de prospectos que son susceptibles de campañar. Una vez ejecutados estos modelos, el sistema entregará un informe al analista con el resultado de la ejecución, donde se incluirán variables de desempeño (como por ejemplo, tiempo de ejecución, número de registros procesados, etc.) así como también variables de resumen (como por ejemplo, histograma de la distribución de prospectos de acuerdo a clase y deuda), informe que el analista validará y comprobará de acuerdo a criterios definidos previamente. Otra salida de esta ejecución es el listado de clientes con ofertas calculadas a través de los modelos. Si la ejecución es correcta, el ejecutivo enviará el mensaje para que se ejecute el proceso siguiente, que es definir acciones de marketing. En caso contrario, resolverá los problemas que se puedan presentar en la ejecución. En caso de error en las tareas automáticas, se definirán mensajes de error personalizados que entregarán luces respecto de cómo solucionar estos errores eventuales.

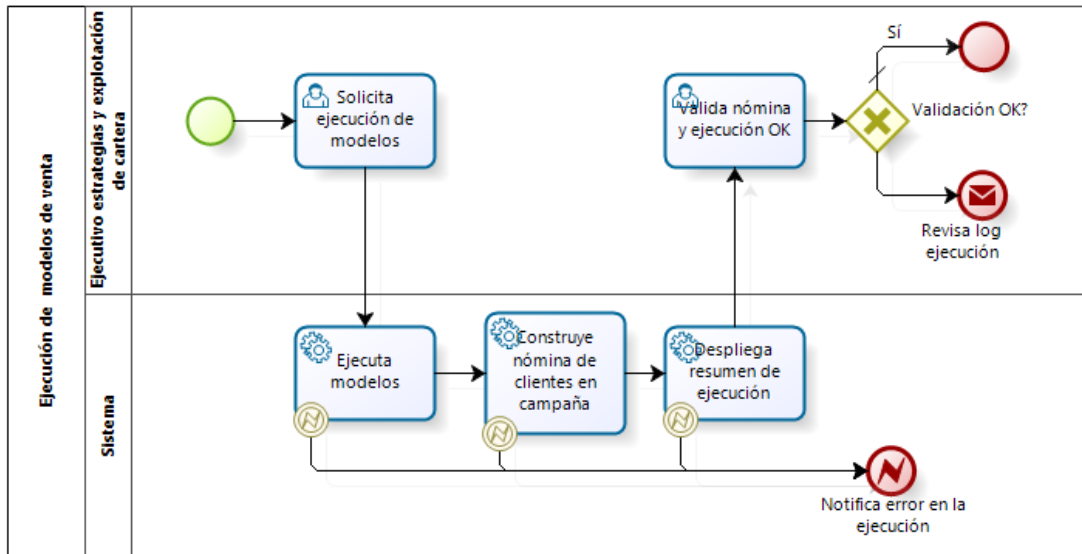


Imagen 19: Diagrama de pistas para Ejecución de modelos de venta

Para simular la ejecución de este diagrama, se utilizaron los siguientes parámetros:







Nº	Componente	Tipo de tarea	Duración/frecuencia de la tarea
1	 Solicita ejecución de modelos	De usuario	Duración distribuida uniformemente entre 1 y 5 minutos.
2	 Ejecuta modelos	De servicio	Duración distribuida normalmente entre 4 horas y 4 horas con 40 minutos. Tiene un evento de error adjunto, el cual ocurre según una distribución gamma con parámetros $k = 1$ y $\theta = 5$.
3	 Construye nómina de clientes en campaña	De servicio	Duración distribuida uniformemente entre 5 y 10 minutos. Tiene un evento de error adjunto, el cual ocurre según una distribución gamma con parámetros $k = 1$ y $\theta = 5$.
4	 Despliega resumen de ejecución	De servicio	Duración distribuida uniformemente entre 1 y 5 minutos. Tiene un evento de error adjunto, el cual ocurre según una distribución gamma con parámetros $k = 1$ y $\theta = 5$.
5	 Valida nómina y ejecución OK	De usuario	Duración distribuida uniformemente entre 30 minutos y 2 horas.
6	 Validación OK?	Puerta exclusiva	Frecuencia distribuida como: 90% de los casos ejecutados sin reparos, 10% de los casos ejecutados erróneamente.

Tabla 16: Parámetros configurados para la simulación de Ejecución de modelos de venta

Este proceso también se simuló en Bizagi, obteniendo los resultados mostrados en la Tabla 17:

Nombre	Tipo	Instancias completadas	Instancias iniciadas	Tiempo mínimo (m)	Tiempo máximo (m)	Tiempo prom. (m)	Tiempo total (m)
Ejecución de modelos de venta	Proceso	1	2	318.59	318.59	318.59	318.59
NoneStart	Evento de inicio	2					
Solicita ejecución de modelos	Tarea	1	2	4.71	4.71	4.71	4.71
Ejecuta modelos	Tarea	1	1	242.41	242.41	242.41	242.41
Construye nómina de clientes en campaña	Tarea	1	1	9.55	9.55	9.55	9.55
Despliega resumen de ejecución	Tarea	1	1	3.36	3.36	3.36	3.36
Notifica error en la ejecución		0					
Valida nómina y ejecución OK	Tarea	1	1	58.55	58.55	58.55	58.55
Validación OK?	Tarea	1	1				
NoneEnd	Compuerta	1					
Revisa log ejecución	Evento de Fin	0					

Tabla 17: Resultados de la simulación de Ejecución de modelos de venta

4 CONSTRUCCIÓN DE LOS MODELOS DE VENTA

Como se ha venido adelantando previamente, la metodología KDD es la utilizada en el desarrollo de estos modelos. A continuación describiremos cada uno de los pasos aplicados al presente proyecto.

4.1 El problema

Lo que se desea modelar es el gasto en riesgo de los clientes, 12 meses después del otorgamiento de nuevos créditos de consumo. Se trata de un problema supervisado, donde se trata de predecir una variable objetivo. Sin embargo, previo a la construcción, se segmenta la cartera de forma de encontrar *clusters* que hagan más precisas las predicciones.

4.2 Definición de la variable objetivo

Al tratarse de un modelo supervisado, existe una variable objetivo que se quiere predecir. Esta variable objetivo equivale al gasto en riesgo que el cliente generará 12 meses después de tomada la colocación. El gasto en riesgo se mide como:

$$GR_{12} = Provisión_{t=12} + \sum_{t=1}^{12} Castigo_t$$

Ecuación 1: Gasto en riesgo luego de 12 meses de colocado un crédito

Al colocar un crédito, se deben constituir provisiones de crédito por él. Cuando un crédito entra en una situación de 180 días de impago o más, parte de él o la totalidad pasan a constituirse como pérdida o castigo.

De acuerdo al problema planteado, lo que se quiere es predecir el gasto en riesgo alcanzado en 12 meses, pero no interesa conocer este monto como valor absoluto (en pesos), sino que lo que interesa es conocer un ratio sobre el saldo de la colocación al mes 12 (tanto en cartera vigente como castigada). No se utiliza la colocación inicial, debido a que pueden existir distorsiones por posibles prepagos anticipados o renegociaciones. La variable objetivo entonces viene a ser la Tasa de Gasto en 12 meses definida por:

$$TG_{12} = \frac{Provision_{t=12} + \sum_{t=1}^{12} Castigo_t}{Colocación_{t=12} + \sum_{t=1}^{12} Castigo_t}$$

Ecuación 2: Variable objetivo

Esta puede tomar valores mayores a cero y menores o iguales a 1.

Por ejemplo, si un crédito cuyo saldo inicial era de \$1.000.000 y que fue pactado a 48 meses a 1.3% de interés mensual, tendrá 48 cuotas mensuales iguales de \$28.136. Supongamos que este crédito se canceló al día durante los primeros 4 meses, donde se constituyó una provisión equivalente al 1% del saldo del crédito¹² mensualmente. Sin embargo, al quinto mes el crédito dejó de cancelarse, entrando en una situación de mora definitiva, aumentando el porcentaje de provisión al 7% el quinto mes, 10% el sexto mes y 50% el séptimo mes. Tendremos entonces el siguiente cuadro de pagos y provisiones:

Mes	Saldo	PI	Provisión	Castigo	GR	Situación contable
1	\$ 1,000,000	1%	\$ 10,000	\$ -	\$ 10,000	Pagado
2	\$ 979,167	1%	\$ 9,792	\$ -	\$ 9,792	Pagado
3	\$ 958,333	1%	\$ 9,583	\$ -	\$ 9,583	Pagado
4	\$ 937,500	1%	\$ 9,375	\$ -	\$ 9,375	Pagado
5	\$ 916,667	7%	\$ 64,167	\$ -	\$ 64,167	Mora 1 a 30 días
6	\$ 916,667	10%	\$ 91,667	\$ -	\$ 91,667	Mora 31 a 60 días
7	\$ 916,667	50%	\$ 458,333	\$ -	\$ 458,333	Mora 61 a 90 días
8	\$ 916,667	50%	\$ 458,333	\$ -	\$ 458,333	Mora 91 a 120 días
9	\$ 916,667	50%	\$ 458,333	\$ -	\$ 458,333	Mora 121 a 150 días
10	\$ 916,667	50%	\$ 458,333	\$ -	\$ 458,333	Mora 151 a 180 días
11	\$ -	-	\$ -	\$ 916,667	\$ 916,667	Castigo
12	\$ -	-	\$ -	\$ 916,667	\$ 916,667	Castigo

Tabla 18: Cuadro de pago de ejemplo de crédito que entra en castigo

En este ejemplo, el crédito se reconoce como 100% castigado al onceavo mes de madurez. Los créditos castigados se deben seguir informando como pérdidas, por un máximo de 3 años de historia de acuerdo a regulaciones vigentes de la SBIF. La TG_{12} en este ejemplo toma el valor de 100%.

4.3 Datos disponibles

Bci cuenta con un *data ware house* corporativo. Se trata de un repositorio de bases de datos enorme, con más de 3.000 vistas y tablas con datos des normalizados en distinto nivel de agrupación, algunos agrupados por cliente, otros agrupados por operación. La mayoría de estas tablas y vistas contiene información de al menos 10 años de historia, lo que dificulta la extracción y consolidación de datos. Sin embargo, la Gerencia de Riesgo cuenta con un *data mart* interno que cuenta con cubos de información más estructurados de acuerdo a lo que se

¹² Este porcentaje se denomina *Probabilidad de incumplimiento* (PI) y se calcula mensualmente para cada cliente en base a modelos estadísticos. Sin embargo, es una medida del default para el stock de la cartera, no para el flujo.

necesita, con información consolidada y agrupada a los niveles que se requieren para la construcción de los modelos.

Los datos disponibles podemos agruparlos en las siguientes familias de datos:

- Datos demográficos: tales como edad, nivel de ingresos, sexo, nacionalidad.
- Datos de perfil: tales como tenencia y antigüedad de cuenta corriente, score de riesgo, banca, sucursal donde opera, tenencia de abono de remuneraciones, tenencia de filtros de riesgo, indicador de si el cliente es trabajador de Bci.
- Datos comportamentales: tales como nivel y evolución histórica de deudas, instituciones en las que mantiene deuda, uso de las líneas disponibles, probabilidad de incumplimiento, tenencia de distintos productos de crédito al interior de Bci, tenencia de créditos renegociados, tenencia de créditos en cartera vencida o castigados, porcentaje de deuda que el cliente mantiene en Bci en relación al sistema financiero (SOW), monto de deuda en relación a la renta (*leverage*) medido antes y después del *course*.
- Datos de *courses*: tales como tipo de operación cursada, monto de la colocación, fecha de *course*, valor de la UF al momento de *course*, precio cobrado, *spread*.

4.4 Selección de registros

Para el desarrollo de los modelos, existen algunos filtros que se aplican para la selección de casos a analizar. Estos son:

- El periodo de análisis comprende *courses* de créditos desde enero 2012 hasta septiembre 2013, de manera de tener 21 camadas mensuales con una madurez mínima de 12 meses cada una (se cuenta con información de desempeño por lo menos hasta septiembre de 2014).
- Se consideran sólo nuevos créditos de consumo, es decir, no se consideran renegociaciones.
- Se selecciona sólo a clientes del Banco *Retail*, excluyendo otro tipo de clientes (por ejemplo, clientes del Banco *Empresarios* que cursan créditos de consumo).

4.5 Segmentación de la base

Previo al modelamiento, se establece la necesidad de segmentar la base para mejorar la predicción de la variable objetivo. Esta segmentación se realiza mezclando criterios de negocio con criterios estadísticos. Se establece la necesidad (desde el punto de vista del negocio) de dividir la base entre colaboradores y no colaboradores, ya que para ambos segmentos las políticas de riesgo son muy distintas.

En relación a los no colaboradores, se aplicó la técnica de K-Medias para determinar los centroides de hasta 4 *clusters*. La restricción del negocio es que estos segmentos sean pocos, fáciles de identificar y replicables en el tiempo.

Los cuatro segmentos identificados entonces son:

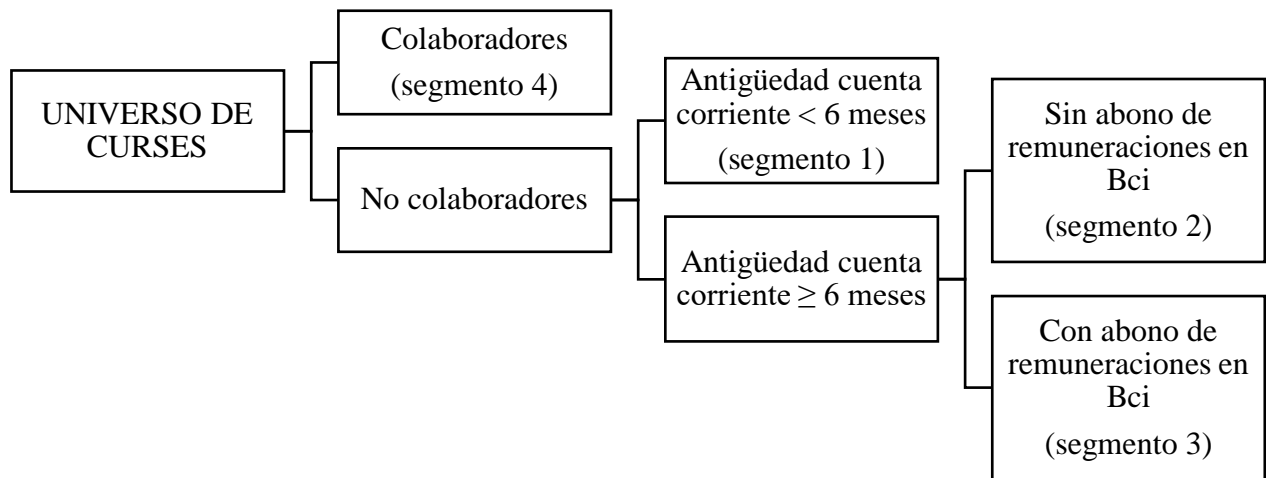


Imagen 20: Segmentación de la base

4.6 Modelamiento

Dentro de las técnicas disponibles para predecir la variable objetivo se encuentran las siguientes como principales.

4.6.1 Regresión lineal

La regresión lineal es una técnica de predicción de una variable continua dentro de un dominio, que busca disminuir el error cuadrático de la desviación entre la predicción y el valor real de la variable objetivo. Siendo y_i el valor que toma la variable objetivo para el registro i , y siendo X_i un vector de características que tratan de predecir el valor de y_i , la regresión lineal, la función que describe el valor de esta variable objetivo es:

$$y_i = \beta X_i + \varepsilon_i$$

Ecuación 3: Regresión lineal

Donde ε_i representa el error de predicción. La solución de este problema consiste entonces en encontrar los valores del vector β , para replicar en otras observaciones referentes al mismo problema.

Un ejemplo gráfico de este problema se puede ver en la Imagen 21, donde los puntos representan los valores observados en una muestra y la recta es la función que predice el valor de cada combinación (X_i, y_i) . En este ejemplo, el vector X_i contiene sólo una característica, y la pendiente de la recta representa el valor de β .

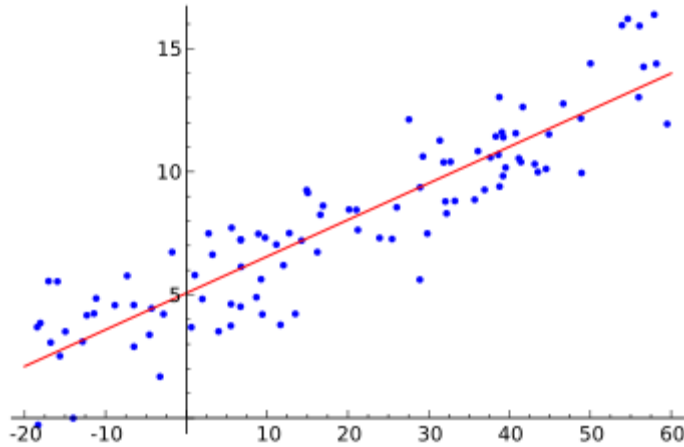


Imagen 21: Ejemplo de una regresión lineal

Para que un problema se pueda modelar a través de este método, se requiere que existan una serie de restricciones:

- La relación entre las características de X_i con y_i debe ser lineal.
- El error debe distribuirse con media 0 y varianza constante (homocedasticidad).
- Las características que componen el vector X_i no deben estar correlacionadas entre sí.

En el problema que estamos tratando de modelar no necesariamente se cumplen estos supuestos. Para poder utilizar esta técnica, se requiere aplicar una transformación previa a los datos, para introducir linealidad en el modelo. Esto conlleva a que otros modelos lineales generalizados tampoco sean aplicables sin estas transformaciones previas.

4.6.2 Estimación por máxima verosimilitud

La estimación por máxima verosimilitud consiste en establecer una función de verosimilitud que establezca los valores de los parámetros que expliquen la variable objetivo. Sea f la función de distribución de las k características X_i que se tienen para la observación i , y $\theta(X_i)$ una función que explica la distribución de X_i . Al asumir independencia entre las k características que componen el vector X_i , se tiene que la función de verosimilitud es:

$$\mathcal{L}(\theta|X_i) = \prod_{j=1}^k f(x_j|\theta)$$

Ecuación 4: Función de verosimilitud

Al maximizar esta función, se despejan los valores de θ_j que explican y_i .

Esta técnica, si bien estima los valores de los parámetros de forma relativamente certera, es algo difícil de explicar y su aplicación se vuelve un poco complicada.

4.6.3 Redes neuronales

Quizás es una de las técnicas más populares de los últimos años, debido a que se moldea muy bien a los datos disponibles para predecir la variable objetivo. Representan un aprendizaje artificial de un problema en el cual se toman variables de entrada (denominadas neuronas en la capa de entrada), se correlacionan entre sí en una capa oculta, se asignan pesos relativos a cada relación entre neuronas y se obtiene como resultado una capa de salida con la predicción, de acuerdo a lo aprendido en cada capa de la red. La técnica de redes neuronales entrena el modelo muchas veces, para determinar los pesos óptimos de las relaciones entre cada par de neuronas. La Imagen 22 muestra conceptualmente cuál es la definición de una red neuronal.

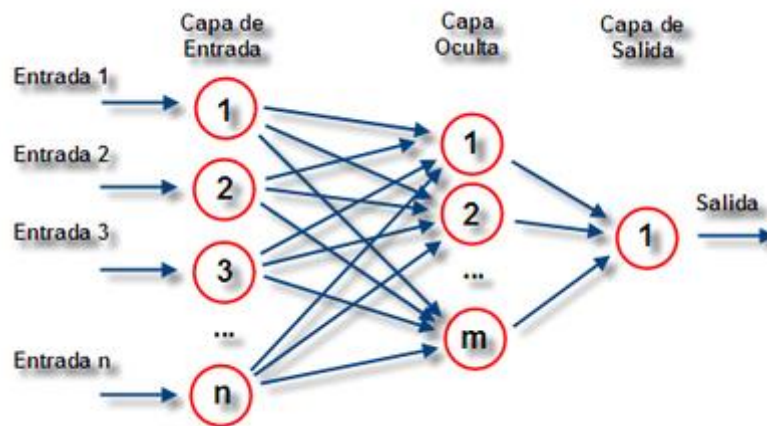


Imagen 22: Definición conceptual de una red neuronal

Si bien los resultados prácticos que se obtienen utilizando este tipo de modelos son más precisos, cuentan con dos principales falencias: tienden a sobre ajustar los modelos (el modelo cuenta con muchos grados de libertad), prediciendo muy bien con los datos de entrenamiento, pero con un desempeño menor en un conjunto de datos diferente (la red “se entrena demasiado bien”).

Adicionalmente, la complejidad al interior de la misma red hace difícil su auditoría por parte de terceros y su replicación e interpretación.

4.6.4 Support Vector Machines (SVM) y Support Vector Regression (SVR)

La técnica de SVM sirve para problemas de clasificación, y consiste en dividir el hiperplano entre las distintas clases, maximizando la distancia entre las distintas clases y minimizando el error de clasificación. Para entenderlo mejor, podemos ver en la Imagen 23 un ejemplo de ello. Acá contamos con dos variables de entrada, por simplicidad: ingreso y edad de un cliente. Lo que se pretende es discriminar entre clases (puntos rojos y puntos azules). En este ejemplo, ambas clases están completamente separadas, y la técnica de SVM determina el hiperplano separador (en este caso, la línea denominada Vector clasificador). Este vector clasificador tiene las dos propiedades ya mencionadas: maximiza la separación entre clases y minimiza el error en la predicción.

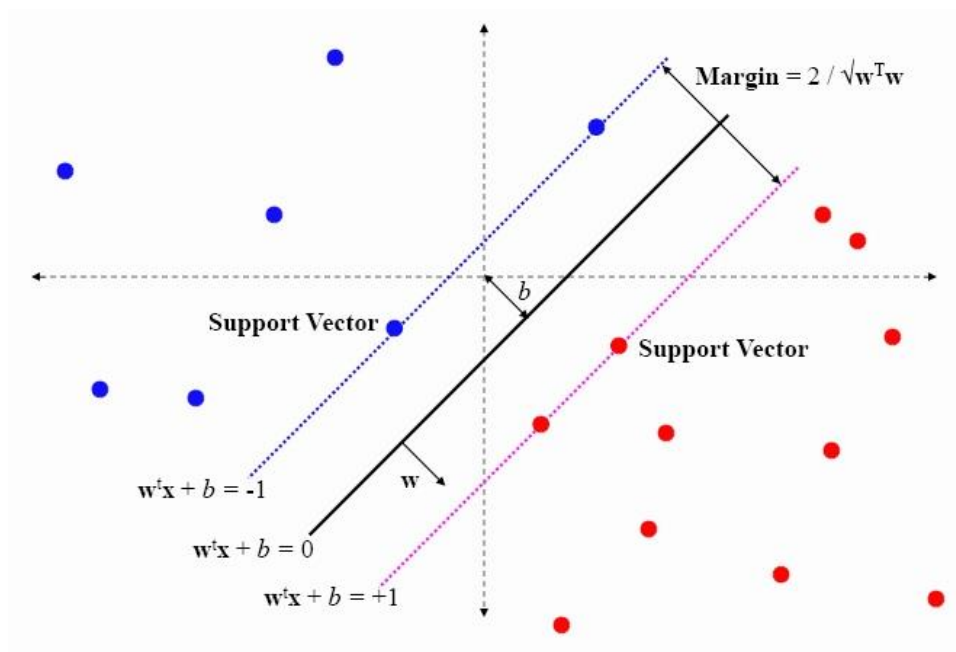


Imagen 23: Definición conceptual de la técnica SVM

Una variación de SVM es SVR, donde la salida del modelo ya no es una clase, sino un valor continuo.

4.6.5 Construcción de una *scorecard* utilizando regresión logística

Una regresión logística es una función de clasificación que asigna distintos pesos a cada variable en el modelo para entregar como resultado la probabilidad de pertenencia a una clase en particular. Se caracteriza por la función:

$$y_i = \frac{1}{1 + e^{-f(x_i)}}$$

Ecuación 5: Regresión logística

Donde $f(X_i)$ es una función de enlace que se estima a través de una regresión lineal, y se denomina logit. En la Imagen 24 podemos ver un ejemplo práctico de este método

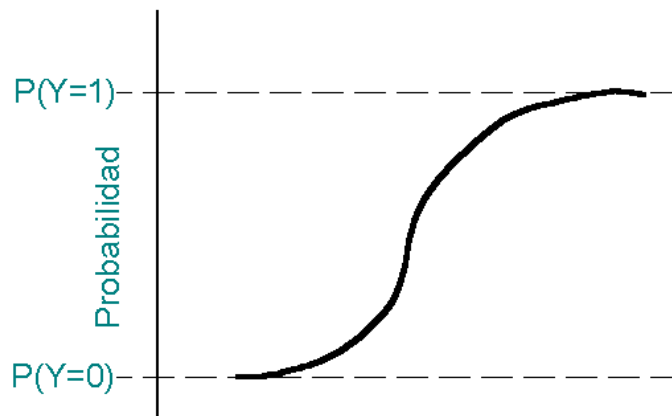


Imagen 24: Ejemplo de una regresión logística

La regresión logística en sí entrega valores discretos (clases), pero se puede adaptar para introducir continuidad en la predicción, tal como veremos más adelante. La idea principal es crear tramos de las variables de entrada X , de manera de contar con K variables dicotómicas (que toman valores 1 ó 0) y determinar los β asociados a cada una de ella. Los β resultantes luego se transforman para crear puntajes asociados a cada tramo de las variables de entrada X , lo que se traduce en la aplicación de una *scorecard*. Dado que esta fue la técnica utilizada finalmente en la construcción de los modelos para 3 de los 4 segmentos detectados, detallaremos con mayor detalle esta técnica a continuación.

La regresión logística en sí necesita una variable dicotómica a predecir. Dado que no se conoce *ex ante* esta variable, lo que se hace en la práctica es duplicar artificialmente la muestra, de

manera que un mismo caso esté marcado con la clase 0 y su duplicado esté marcado con la clase 1. Luego, cada registro se pondera por una medida relativa de la colocación del cliente sobre el total de colocación de la clase, y por la TG_{12} . Así, se construye una variable w_i que tomará los siguientes valores:

$$w_i = \begin{cases} TG_{12,i} \times \frac{Colocación_i}{\sum_{i=1}^N Colocación_i}, & \text{si } clase_i = 1 \\ (1 - TG_{12,i}) \times \frac{Colocación_i}{\sum_{i=1}^N Colocación_i}, & \text{si } clase_i = 0 \end{cases}$$

Ecuación 6: Construcción de la variable “peso”

Si la TG_{12} para un cliente en particular es de 5%, al ser clasificado en la clase 1 tendrá menor peso que su clon de la clase 0. En cambio, si la TG_{12} toma un valor más elevado, por ejemplo 70%, cuando pertenezca a la clase 1 tendrá un peso relativamente menor que cuando pertenezca a la clase 0. Esto quiere decir que, cuando un cliente es relativamente “bueno” (es decir, tiene una TG_{12} relativamente baja) o cuando sea relativamente “malo” (cuando tenga una TG_{12} relativamente alta) tendrá más peso en la muestra. Si la muestra tiene un desbalanceo a favor de los clientes con menor TG_{12} (como es el caso de la muestra de desarrollo), este peso tiende a balancear mejor las observaciones. Esto implica que no importa cuál es el punto de corte de la variable objetivo para clasificar a un cliente como “bueno” o “malo”, ya que la técnica permite una construcción artificial de esta clasificación en torno a una variable objetivo continua. Finalmente, como un cliente con una colocación mayor impacta de mayor forma que un cliente con una colocación menor, w_i ayuda también a corregir por el efecto de tener una menor o mayor colocación.

En este proyecto se decidió utilizar esta técnica. Esta decisión se fundamenta por sus beneficios, tal como se describe en Rajan y Tulasi (2015), los cuales son:

- Es de fácil entendimiento, lo que facilita también la explicación a los aprobadores y auditores.
- Es de fácil programación en casi cualquier sistema que pueda hacer cálculos simples, como la suma.
- No importa la escala de las variables, lo que importa es su tramificación.
- No importa que los atributos estén correlacionados entre sí.
- No es necesario que exista una variable objetivo dicotómica.
- No importa que se tenga una muestra desbalanceada, ya que se balancea por el peso relativo.
- El peso relativo ayuda a dar más importancia a clientes con mayor colocación, que impactan de mayor forma en el nivel de GR_{12} .

4.7 Transformación de los atributos

Las características de la base o atributos presentan distintas escalas y distintos valores que pueden distorsionar los resultados. Por ejemplo, existen variables con un rango relativamente acotado como la edad, que parte desde los 18 a los 77 años o la probabilidad de incumplimiento, que va de 0 a 1, y variables con un rango mucho más amplio como la deuda vigente en el sistema financiero, que van desde los \$0 a valores bastante elevados (sobre los \$1,400,000,000).

Para ser consistente con la técnica de la *scorecard*, es necesario entonces transformar los atributos de entrada. Existen diversas técnicas para realizar esto, tales como la estandarización (que consiste en aplicar una transformación corrigiendo por la media y varianza de la distribución de la variable en la muestra), aplicar transformaciones logarítmicas (para suavizar la distribución de las variables), escalamiento (que consiste en asignar valores a los atributos en un rango determinado, por ejemplo, entre 0 y 1), entre otras. La que se aplicó en el desarrollo de este modelo es la de tramificación. A cada atributo correlacionado con la variable objetivo se le determinan intervalos óptimos donde se puedan establecer distintos tramos, y cada uno de estos tramos luego se convierte en un nuevo atributo (que toma el valor de 1 en caso de pertenecer a los valores establecidos por el intervalo, y 0 en otro caso).

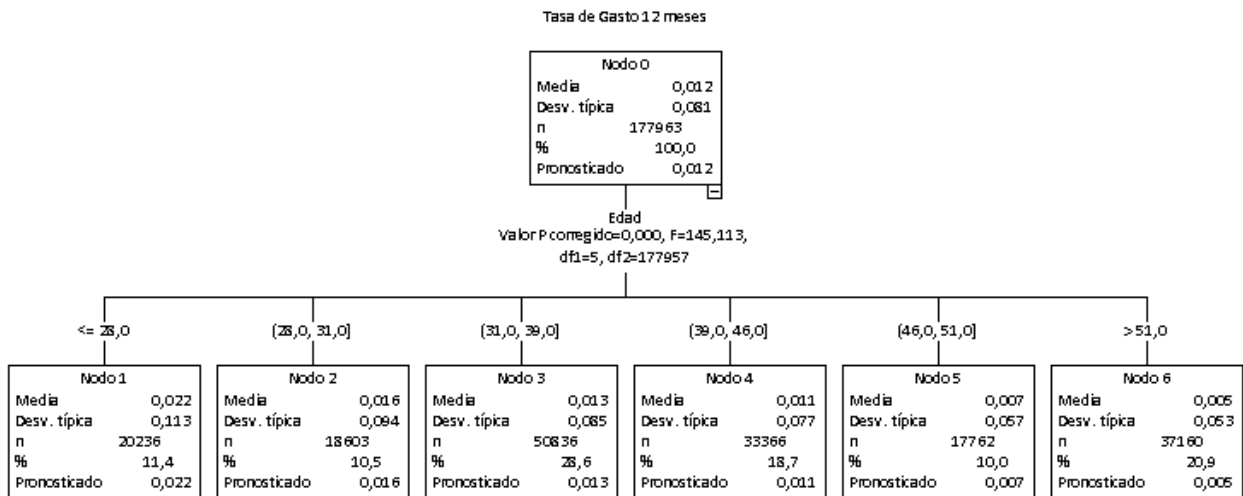


Imagen 25: Ejemplo de tramificación de la variable Edad

En la Imagen 25 podemos ver la transformación aplicada sobre la variable Edad. Esta variable está correlacionada con la variable objetivo (TG_{12}), y a través de un árbol de clasificación se determinan los distintos intervalos que puede tomar para construir, en este caso 6 nuevas variables ($Edad1_i = \{1 \text{ si } Edad \leq 28\}$, hasta $Edad6_i = \{1 \text{ si } Edad > 51\}$). Esta transformación se aplica para cada segmento identificado, para cada conjunto de atributos correlacionados en cada segmento con la variable objetivo. Para saber si existe correlación entre

las variables se utilizó los coeficientes de correlación de Pearson para variables discretas y de Spearman para variables continuas. Como única restricción, se necesita que dentro de las variables independientes se considere el *leverage*¹³ final alcanzado por los clientes. Si bien se cuenta con esta variable al momento de desarrollar los modelos, esta variable no se tendrá disponible al momento de aplicar el modelo a observaciones nuevas, pero tiene la particularidad de que servirá para establecer los escenarios de posibilidades de un cliente, y con ello poder utilizar esta variable como una variable de decisión en la aplicación de un modelo de endeudamiento óptimo.

4.8 Solución del modelo

Tal como se adelantó previamente, se identificaron 4 segmentos de clientes. Por el momento dejaremos de lado el segmento de colaboradores, el que será visto en el punto 4.8.3. Veremos en detalle la solución de cada uno de los otros tres segmentos. Cabe mencionar que para el desarrollo se trabajó con el 75% de los datos, y se reservó un 25% para validar el modelo.

4.8.1 Segmento 1: clientes con antigüedad < 6 meses (clientes nuevos)

Las variables candidatas del modelo son:

Variable	Núm. casos	Mínimo	Máximo	Media	Desviación estándar
Edad	24,107	19	79	35.50	9.33
Renta líquida (M\$)	24,107	0	256,665	1,083.7	1,829.49
PI	24,107	0%	88.37%	1.26%	2.61%
Antigüedad (meses)	24,107	0	5	2.53	1.61
Leverage previo al curse	23,236	-1.00	173.47	3.37	4.35
Leverage post curse	24,107	0	126,686.00	73.11	1,502.91
Núm. inst. financieras	23,224	0	7	1.54	1.02
Lín. disp. sobre rta. previo al curse	23,228	0	47,242.00	27.28	593.57
Score de riesgo (R04)	24,107	113	283	202.67	38.81
Tenencia de hipotecario	24,107	0	1		
N válido (por lista)	23,224				

Tabla 19: Variables candidatas para el segmento 1

Cabe mencionar que, de todas las variables descritas en 4.3, estas son las que hacen sentido al negocio y se correlacionan con la variable objetivo. Dado el bajo número de valores perdidos, se decide eliminar estos registros. A continuación una breve explicación de estas variables.

¹³ El *leverage* o nivel de endeudamiento de un cliente corresponde a la deuda de corto plazo de un cliente, sobre su renta. Indica un nivel de apalancamiento de los pasivos de un cliente en relación a sus ingresos.

- Edad: corresponde a la edad del cliente al momento de cursar la operación.

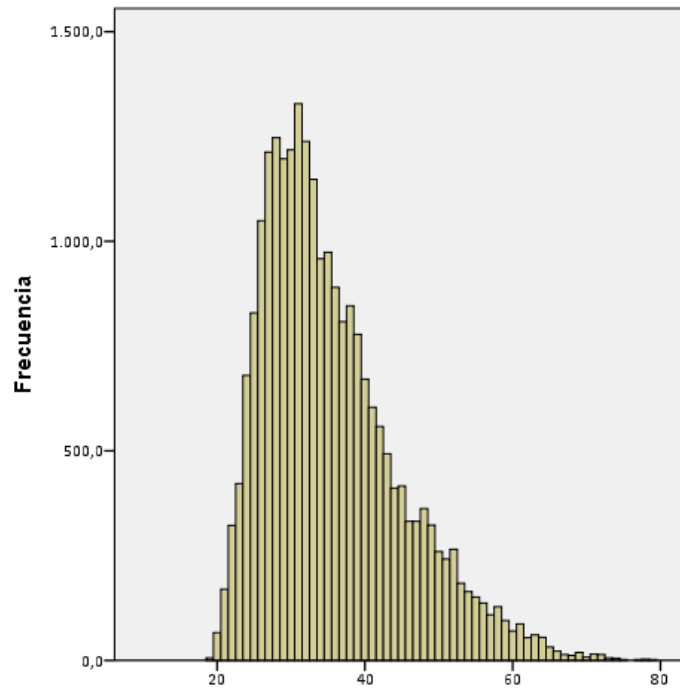


Imagen 26: Histograma de la variable Edad

- Renta líquida: es la suma de sus ingresos acreditados durante el mes. Existen valores extremos y valores fuera de rango, para lo cual se aplica una transformación adicional. Estos valores se truncan a \$10,000,000, valor que es el máximo nivel de renta líquida mensual atendible en el banco *Retail*. Como no son muchos casos, la distribución no se distorsiona, tal como se puede ver en la Imagen 28.

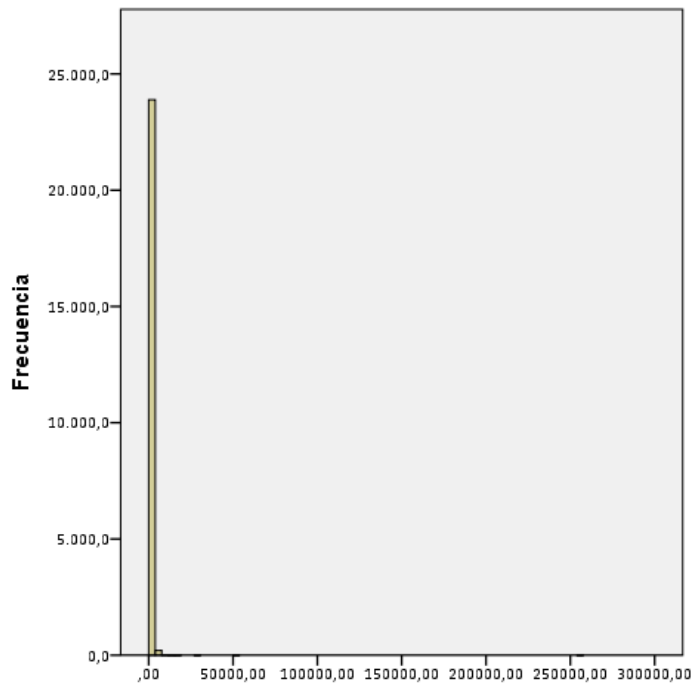


Imagen 27: Histograma de la variable Renta Líquida antes de la transformación

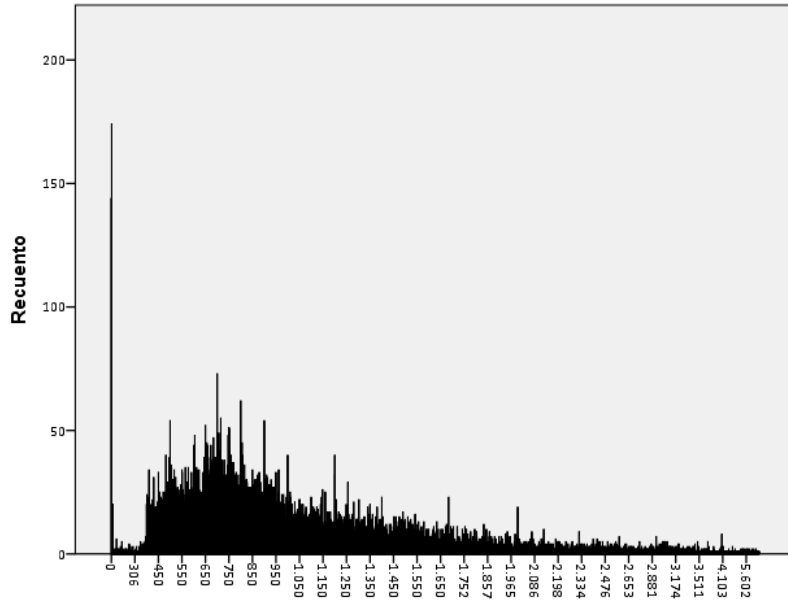


Imagen 28: Histograma de la variable Renta Líquida después de la transformación

- PI: corresponde a la probabilidad de incumplimiento, determinada por modelos estadísticos que recogen variables de comportamiento de la cartera. Miden la proporción de gasto por provisiones de crédito que el banco debe hacer mensualmente. Se mide

como porcentaje, y toma valores mayores a cero y menores o iguales a 100%. Corresponde a un *score* comportamental de la cartera, que no tiene relación necesariamente con la originación de nuevos créditos.

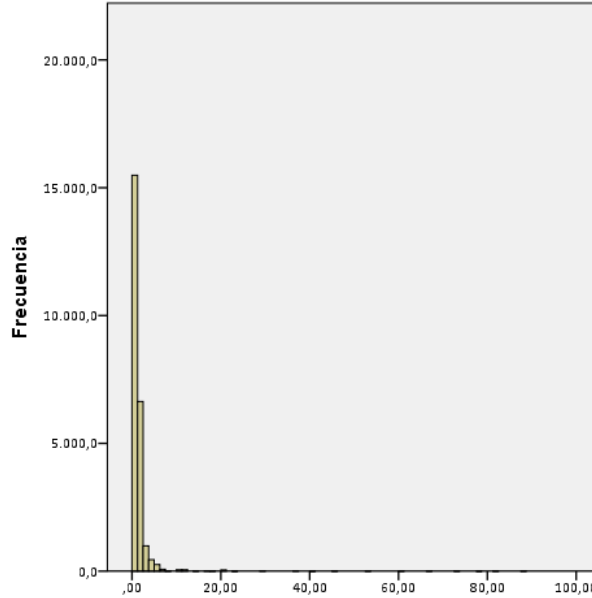


Imagen 29: Histograma de la variable Probabilidad de Incumplimiento (en porcentajes, de 0 a 100)

- Antigüedad en cuenta corriente del cliente al momento de cursar el crédito, medida en meses. Dado que este segmento específicamente considera clientes con menos de 6 meses de antigüedad, se tendrá una distribución truncada de los datos.

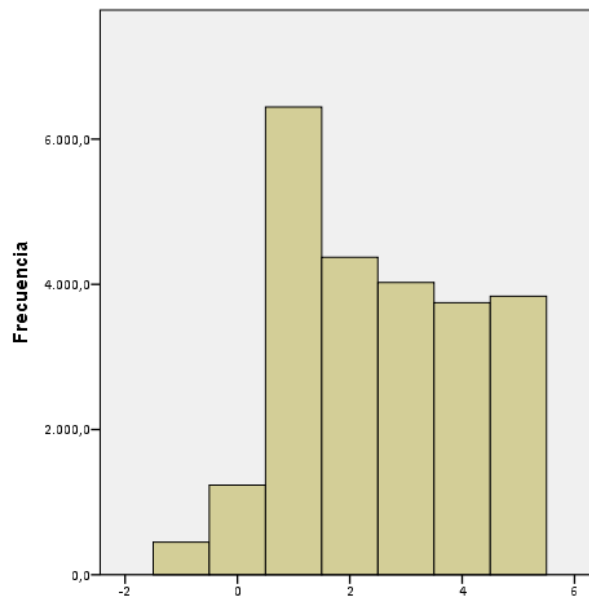


Imagen 30: Histograma de la variable Antigüedad en Cuenta Corriente

- *Leverage* previo al curse: corresponde al nivel de endeudamiento en el sistema financiero que tiene el cliente previo a tomar el crédito, medido como el total de la deuda de consumo sobre la renta del cliente.

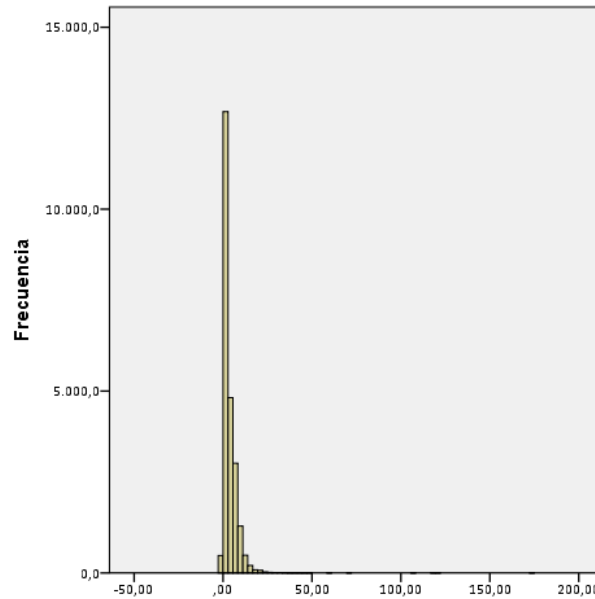


Imagen 31: Histograma de la variable Leverage previo al curse

- *Leverage* post curse: corresponde al nivel de endeudamiento en el sistema financiero que tiene el cliente dos meses después de tomar el crédito, medido como el total de la deuda de consumo sobre la renta del cliente (equivale a “las veces renta” de deuda que mantiene un cliente). Como se adelantó previamente, esta variable no estará disponible cuando se aplique el modelo en datos nuevos. Sin embargo, la utilidad de esta variable en el modelo es crucial, ya que en base al nivel de GR_{12} que se desea alcanzar es que se define el nivel de endeudamiento óptimo de cada cliente, que se traduce posteriormente en una oferta de campaña.

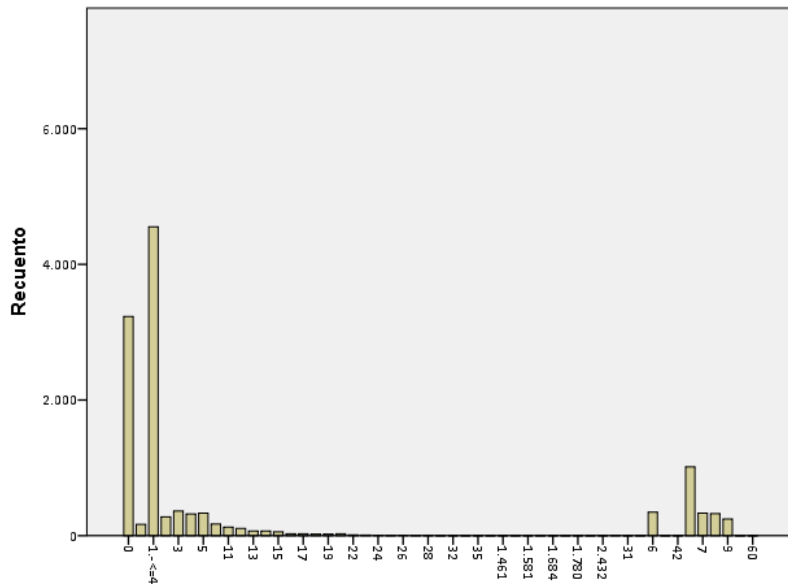


Imagen 32: Histograma de la variable Leverage Post Course (corregido por transformación de renta)

- Número de instituciones financieras (o número de acreedores): indica el número de bancos y otras instituciones financieras donde el cliente mantiene deudas. Cada banco entrega un límite de endeudamiento distinto, pero la evaluación debe hacerse integralmente sobre la capacidad de pago del cliente en todo el sistema financiero.

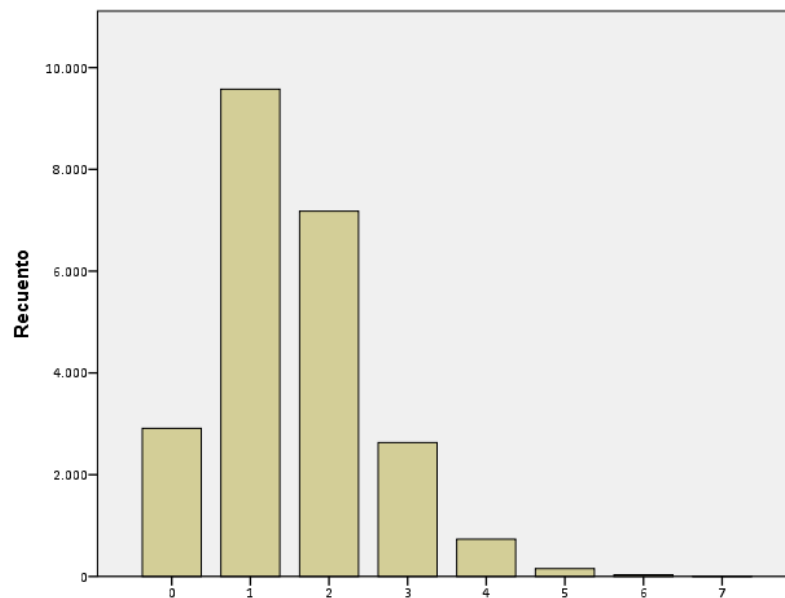


Imagen 33: Histograma de la variable Número de acreedores

- Líneas disponibles sobre renta previo al curse: las líneas disponibles representan el nivel de endeudamiento disponible en todas las instituciones financieras que posee un cliente

y que no ha utilizado aún. Por ejemplo, cuando le otorgan productos de crédito rotativos (líneas de sobregiro o tarjetas de crédito), éstas tienen un límite designado. Si a este límite se le descuenta el uso del cliente en dichos productos, pasa a llamarse líneas disponibles. Es una medida del endeudamiento potencial, medida en relación a la renta del cliente.

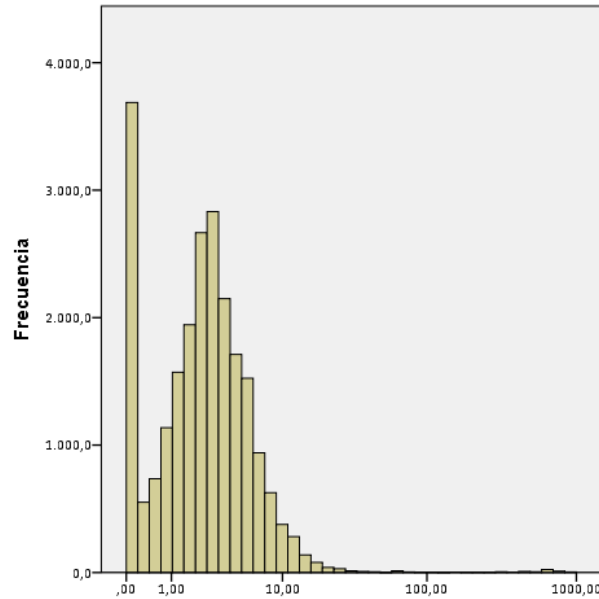


Imagen 34: Líneas disponibles sobre renta previo al curso

- *Score* de riesgo: Bci cuenta con un modelo interno de comportamiento de clientes basado en su historia. Este *score* se construye considerando los últimos 24 meses de historia y su evolución en el tiempo. Busca predecir comportamiento de pago en el sistema financiero a futuro. En general es bastante predictivo para clientes nuevos, pero pierde un poco de poder predictivo (en relación a las otras variables) a medida que aumenta la antigüedad del cliente en Bci. Es un puntaje discreto, que va de 100 a 283 puntos, donde a mayor puntaje mejor es el comportamiento de pago del cliente. Al igual que la PI, es un *score* comportamental, pero a nivel del sistema financiero, y no necesariamente cuenta con las mismas propiedades que un *score* de originación.

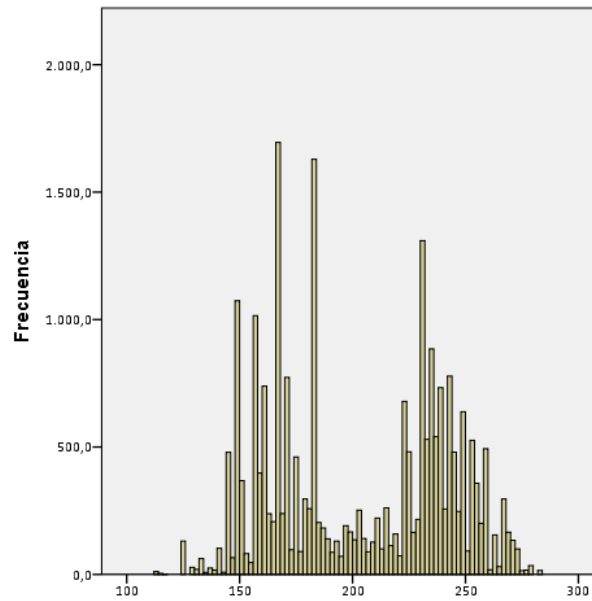


Imagen 35: Histograma de la variable Score de riesgo (R04)

- Tenencia de créditos hipotecarios: la tenencia de un crédito hipotecario implica generalmente un mejor comportamiento de pago, toda vez que el crédito hipotecario mantiene en prenda la vivienda financiada. Esta variable es una variable dicotómica que toma el valor 1 cuando el cliente tiene hipotecario previo al curse y 0 cuando no.

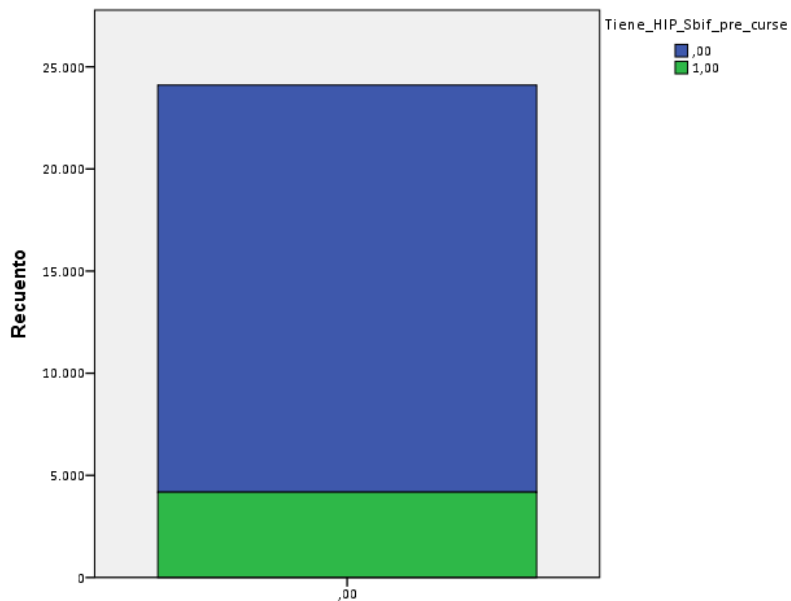


Imagen 36: Distribución de la variable Tenencia de créditos hipotecarios

Como se adelantó, la técnica exige que cada uno de los registros se clone, teniendo uno de ellos el valor de clase 1 y su clon el valor 0. Se asignan luego los pesos relativos para ponderar cada registro y en base a ello, se tramifica cada variable. Para ello, se modela la clase como variable a predecir utilizando cada una de las variables de entrada. A modo de ejemplo, y para simplificar, sólo se presentará el árbol de clasificación de la variable Edad. Esto determina tres nuevas variables (Edad1, Edad2 y Edad3) que toman valores 1 ó 0.

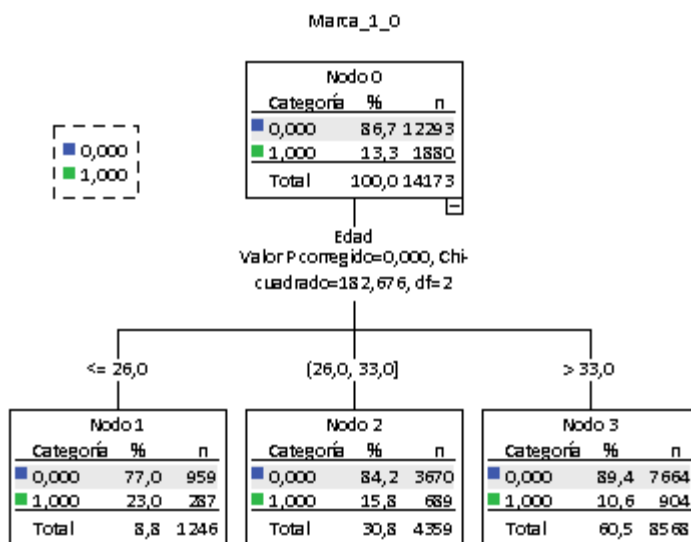


Imagen 37: Árbol de clasificación para determinar los tramos de Edad

Al repetir el procedimiento para todas las otras variables, se obtienen los siguientes tramos¹⁴:

- Edad_1: Edad <= 25
- Edad_2: Edad > 25 and Edad <= 33
- Edad_3: Edad > 33
- PI_1: PIXPDI_New <= 0.260
- PI_2: PIXPDI_New > 0.260 and PIXPDI_New <= 0.817
- PI_3: PIXPDI_New > 0.817 and PIXPDI_New <= 1.414
- PI_4: PIXPDI_New > 1.414
- R04_1: R04 <= 153
- R04_2: R04 > 153 and R04 <= 174
- R04_3: R04 > 174 and R04 <= 229
- R04_4: R04 > 229 and R04 <= 244
- R04_5: R04 > 244
- Lev_pos_1: Leverage > 0 and Leverage <= 4
- Lev_pos_2: Leverage > 4 and Leverage <= 8
- Lev_pos_3: Leverage > 8 and Leverage <= 15
- Lev_pos_4: Leverage > 15
- Lev_pos_5: missing(Leverage) (Sin renta)

¹⁴ Nota: aquí se presentan sólo las variables definitivas del modelo.

- Tiene_HIP_1: Tiene_HIP_Sbif_pre_course = 0
- Tiene_HIP_2: Tiene_HIP_Sbif_pre_course = 1
- Lin_sob_rta_1: Lin_disp_sobre_rta_pre_course <= 0
- Lin_sob_rta_2: Lin_disp_sobre_rta_pre_course > 0 and Lin_disp_sobre_rta_pre_course <= 4
- Lin_sob_rta_3: Lin_disp_sobre_rta_pre_course > 4 and Lin_disp_sobre_rta_pre_course <= 8
- Lin_sob_rta_4: Lin_disp_sobre_rta_pre_course > 8

Luego, se estima la siguiente función logit:

$$L(X_i) = \beta_0 + \beta_1 \text{Edad}_1_i + \beta_2 \text{Edad}_2_i + \beta_3 \text{Edad}_3_i + \dots + \beta_{25} \text{Lin_sob_rta}_4_i$$

Ecuación 7: Función logit del segmento “clientes nuevos”

La función a estimar a través de una regresión logística viene dada por:

$$y_i(X_i) = \frac{1}{1 + e^{-L(X_i)}}$$

Ecuación 8: Función de regresión logística

Que toma los valores de la clase, 1 ó 0. Los valores de los β_j se estiman, y luego hay que aplicar una transformación para traducirlos en un *score*. Para transformar estos β_j se utilizan los odds-ratio, que vienen a ser la proporción de clientes clasificados como 1 sobre los clientes clasificados como 0. En Bci generalmente se utiliza un puntaje base de 600 puntos en la construcción de *scorecards*, y cada punto adicional de β_j incrementa este puntaje base en 80 puntos¹⁵. Esto implica que cada β_j (excepto β_0) se deben multiplicar por un factor fijo.

Para el segmento de clientes nuevos, la *scorecard* construida (con el paso de β_j a *score*) viene dada por la Tabla 20.

Esto implica que a cada cliente, dependiendo de sus atributos, se le calcula un puntaje final, que va de los 435 a los 1,348 puntos.

Lo que queda ahora es traducir este *score* en un valor de TG_{12} . Para ello, nuevamente se vuelve a modelar un árbol de clasificación que tramifique los valores del puntaje en torno al valor de TG_{12} . El resultado de esta nueva clasificación se puede ver en la Tabla 21, donde además se

¹⁵ Esta escala es arbitraria y definida por la Gerencia de Desarrollo de Modelos Estadísticos de Bci, pero se utilizará la misma en todos los segmentos para hacer los puntajes finales comparables entre sí.

muestran otros indicadores (como la TG_{12}). A cada uno de los clientes pertenecientes a estos tramos se le asigna como $\overline{TG_{12}}$ (predicción) el valor de la TG_{12} real promedio del tramo. Es decir, un cliente con 740 puntos tendrá una $\overline{TG_{12}} = 20.54\%$, mismo valor que tendrá un cliente con 780 puntos.

Tramo de los atributos	β	Sig.	Score	Score máximo	Score mínimo
Edad <= 25	0.510	0.000	-64	0	-64
Edad > 25 and Edad <= 33	0.136	0.011	-17		
Edad > 33		0.000	0		
PIxPDI_New <= 0.260	-1.554	0.000	194	194	0
PIxPDI_New > 0.260 and PIxPDI_New <= 0.817	-1.058	0.000	132		
PIxPDI_New > 0.817 and PIxPDI_New <= 1.414	-0.473	0.000	59		
PIxPDI_New > 1.414		0.000	0		
R04 <= 153	2.269	0.000	-284	0	-284
R04 > 153 and R04 <= 174	1.638	0.000	-205		
R04 > 174 and R04 <= 229	1.110	0.000	-139		
R04 > 229 and R04 <= 244	0.339	0.008	-42		
R04 > 244		0.000	0		
Leverage_post > 0 and Leverage_post <= 4	-0.822	0.000	103	103	-95
Leverage_post > 4 and Leverage_post <= 8	-0.467	0.000	58		
Leverage_post > 8 and Leverage_post <= 15	0.204	0.001	-25		
Leverage_post > 15	0.763	0.000	-95		
missing(Leverage_post) = 1		0.000	0		
Tiene_HIP_Sbif_pre_curse = 0	0.521	0.000	-65	0	-65
Tiene_HIP_Sbif_pre_curse = 1			0		
Lin_disp_sobre_rta_pre_curse <= 0	0.861	0.000	-108	0	-108
Lin_disp_sobre_rta_pre_curse > 0 and	0.846	0.000	-106		
Lin_disp_sobre_rta_pre_curse <= 4					
Lin_disp_sobre_rta_pre_curse > 4 and	0.633	0.000	-79		
Lin_disp_sobre_rta_pre_curse <= 8					
Lin_disp_sobre_rta_pre_curse > 8		0.000	0		
Constante	-3.547	0.000	1,051	1,051	1,051

1,348	435
-------	-----

Tabla 20: Resultado de la regresión logística (clientes nuevos)

Tramo puntaje	Núm casos	Promedio ponderado TG_{12} real	Promedio ponderado TG_{12} predicción	Error predicción	Error cuadrático medio
≤ 675	1,332	48.99%	46.74%	1.61%	15.86%
(675; 738]	1,629	30.80%	29.95%	1.50%	13.40%
(738; 795]	2,021	20.54%	20.78%	1.60%	10.63%
(795; 845]	1,785	14.54%	16.75%	0.25%	9.02%
(845; 881]	1,347	12.36%	12.10%	0.36%	6.45%
(881; 930]	1,917	7.72%	6.98%	2.20%	5.07%
(930; 975]	1,464	5.64%	5.57%	0.85%	3.42%
(975; 1,027]	1,738	3.18%	3.62%	1.16%	2.65%
(1,027; 1,096]	1,514	2.78%	3.06%	0.47%	1.91%
$> 1,096$	1,696	1.62%	1.70%	-0.26%	0.79%
Total	16,443	13.13%	14.27%	1.00%	6.86%

Tabla 21: Predicción de la TG_{12} usando el score (clientes nuevos)

Existen varias formas de medir la precisión en la predicción. Generalmente se utiliza el estadígrafo de Kolmogórov-Smirnov (KS), el indicador de Gini o el *Information Value*. Sin embargo, estos indicadores pierden utilidad, toda vez que las clases (0 y 1) fueron construidas artificialmente. De hecho, sólo a modo de ejercicio, el valor de KS obtenido en este modelo fue de 94.6% (el valor de referencia para aceptar un modelo como válido al interior de Bci es de 30%). En consecuencia, para estos modelos se decide utilizar el error cuadrático medio (ECM) como indicador de bondad de ajuste, el que se calcula como:

$$ECM_i = 1/m \times \sqrt{\sum_{i=1}^m (TG_{12} - \overline{TG_{12}})^2}$$

Ecuación 9: Cálculo del ECM

Donde m representa el número de casos que hay en cada tramo de *score*. Inicialmente, se definió como estándar aceptar modelos donde $ECM / \overline{TG_{12}} < 10\%$, aún cuando el estándar interno del banco es de un 15%.

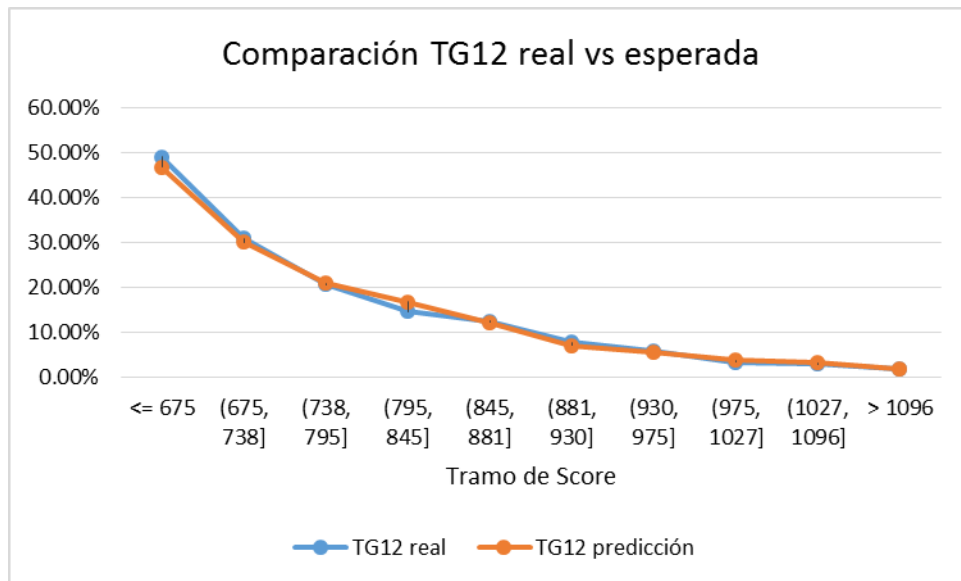


Imagen 38: Comparación gráfica de la TG_{12} por tramo de score (clientes nuevos)

El último paso para aceptar el modelo es el *back-test*, donde al menos se espera lograr estabilidad en el modelo a través del tiempo. Como se puede apreciar en la Imagen 39, la predicción se ajusta bastante bien. En periodos donde se experimenta un alza de la TG_{12} (noviembre y diciembre de 2012) la predicción se mueve en la misma dirección, aunque en un escenario más conservador (la predicción sobre estima el gasto real).

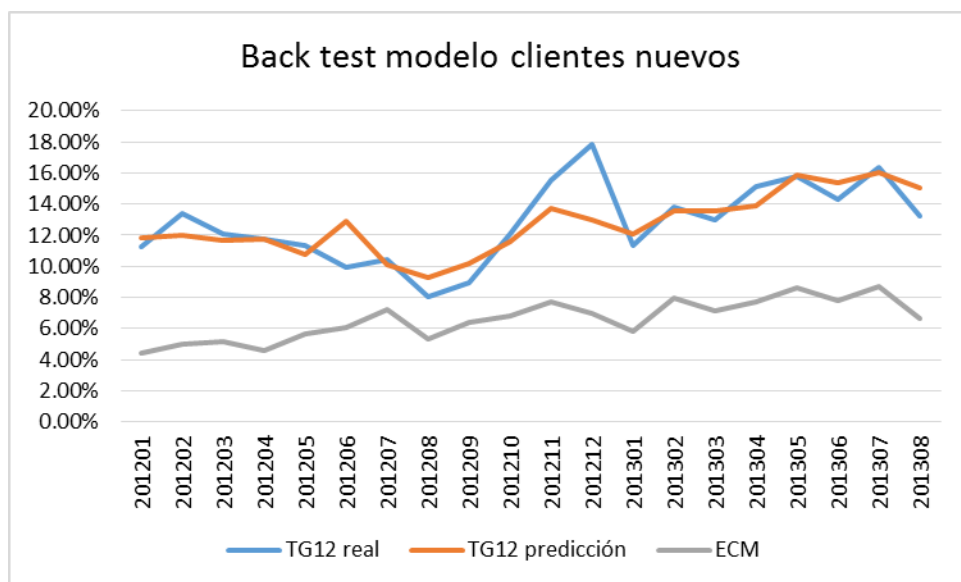


Imagen 39: Estabilidad del modelo a través del tiempo (clientes nuevos)

4.8.2 Segmento 2: clientes con antigüedad ≥ 6 meses y sin abono de remuneraciones en Bci (clientes sin abono de remuneraciones)

El abono de remuneraciones en Bci se constituye cuando la empresa para la cual trabaja el cliente realiza un abono en su cuenta corriente o cuenta a la vista, por concepto de pago de remuneraciones. Esto le trae consecuencias muy positivas a Bci, ya que por lo general los créditos de consumo se cursan contratando el pago automático con cargo en cuenta corriente (PAC), situación que opera de la siguiente forma: cuando cae la fecha de vencimiento de la cuota de un crédito, el sistema del banco va a consultar el saldo en cuenta corriente del cliente. Cuando encuentra saldo suficiente, el sistema realiza el cargo en la cuenta corriente por el monto de la cuota del crédito, y el crédito se paga y se mantiene en una situación contable “al día”. Es mucho más probable encontrar saldo suficiente en la cuenta corriente para realizar este cargo cuando el cliente cuenta con abono de remuneraciones que cuando no lo tiene.

Para simplificar, sólo se mostrará parte de los resultados, ya que la técnica es homóloga al segmento de clientes nuevos.

La solución del modelo se muestra en la Tabla 22. Luego, se reitera el procedimiento de la técnica ya revisada para el segmento de clientes nuevos. Los puntajes calculados individualmente para cada cliente se agrupan en los tramos que se muestran en la Tabla 23.

Tramo de los atributos	β	Sig.	Score	Score máximo	Score mínimo
Edad <= 24	0.911	0.000	-112	0	-112
Edad > 24 and Edad <= 28	0.358	0.000	-44		
Edad > 28 and Edad <= 48	0.221	0.000	-27		
Edad > 48		0.000	0		
PIxPDI_New <= 0.026594	-1.503	0.000	185	185	0
PIxPDI_New > 0.026594 and PIxPDI_New <= 0.1306	-0.989	0.000	122		
PIxPDI_New > 0.1306 and PIxPDI_New <= 0.3635	-0.509	0.000	63		
PIxPDI_New > 0.3635 and PIxPDI_New <= 1.3333	-0.342	0.000	42		
PIxPDI_New > 1.3333		0.000	0		
R04 <= 153	3.168	0.000	-391	0	-391
R04 > 153 and R04 <= 174	2.089	0.000	-258		
R04 > 174 and R04 <= 229	0.875	0.000	-108		
R04 > 229 and R04 <= 244	0.321	0.000	-40		
R04 > 244		0.000	0		
ANTIG_CCT <= 12	0.583	0.000	-72	0	-72
ANTIG_CCT > 12 and ANTIG_CCT <= 24	0.561	0.000	-69		
ANTIG_CCT > 24 and ANTIG_CCT <= 36	0.344	0.000	-42		
ANTIG_CCT > 36		0.000	0		
Leverage > 0 and Leverage <= 4	-1.309	0.000	161	161	0
Leverage > 4 and Leverage <= 8	-1.078	0.000	133		
Leverage > 8 and Leverage <= 12	-0.646	0.000	80		
Leverage > 12 and Leverage <= 15	-0.353	0.000	44		
Leverage > 15 or Leverage <= 0		0.000	0		
Tiene_HIP_Sbif_pre_curse = 0	0.331	0.000	-41	0	-41
Tiene_HIP_Sbif_pre_curse = 1			0		
Lin_disp_sobre_rta_pre_curse <= 1.5	0.495	0.000	-61	0	-61
Lin_disp_sobre_rta_pre_curse > 1.5 and Lin_disp_sobre_rta_pre_curse <= 3	0.333	0.000	-41		
Lin_disp_sobre_rta_pre_curse > 3 and Lin_disp_sobre_rta_pre_curse <= 4.5	0.252	0.000	-31		
Lin_disp_sobre_rta_pre_curse > 4.5 and Lin_disp_sobre_rta_pre_curse <= 6	0.043	0.533 ¹⁶	-5		
Lin_disp_sobre_rta_pre_curse > 6		0.000	0		
Constante	-3.471	0.000	1,053		
				1,399	376

Tabla 22: Resultado de la regresión logística (clientes sin abono de remuneraciones)

¹⁶ Nota: aun cuando la significancia (p-valor) de esta variable para este tramo no presenta significancia estadística, igual se mantiene en la scorecard por requerimiento del negocio.

Tramo puntaje	Núm casos	Promedio ponderado TG_{12} real	Promedio ponderado TG_{12} predicción	Error predicción	Error cuadrático medio
≤ 886	7,772	23.11%	23.26%	-2.03%	10.67%
(886; 959]	6,943	8.12%	8.36%	0.28%	4.21%
(959; 1,008]	6,526	4.87%	4.92%	0.69%	2.89%
(1,008; 1,048]	6,802	3.61%	3.73%	0.12%	1.93%
(1,048; 1,095]	8,256	1.99%	1.99%	0.58%	1.28%
(1,095; 1,117]	4,484	1.62%	1.66%	0.43%	1.02%
(1,117; 1,156]	7,177	1.48%	1.61%	0.15%	0.86%
(1,156; 1,200]	8,430	0.93%	1.01%	0.14%	0.56%
(1,200; 1,259]	9,658	0.74%	0.75%	0.11%	0.43%
$> 1,259$	13,535	0.28%	0.37%	0.00%	0.17%
Total	79,583	4.19%	4.28%	0.02%	2.22%

Tabla 23: Predicción de la TG_{12} usando el score (clientes sin abono de remuneraciones)

Gráficamente, las desviaciones de la predicción de la TG_{12} se pueden ver en la Imagen 40.

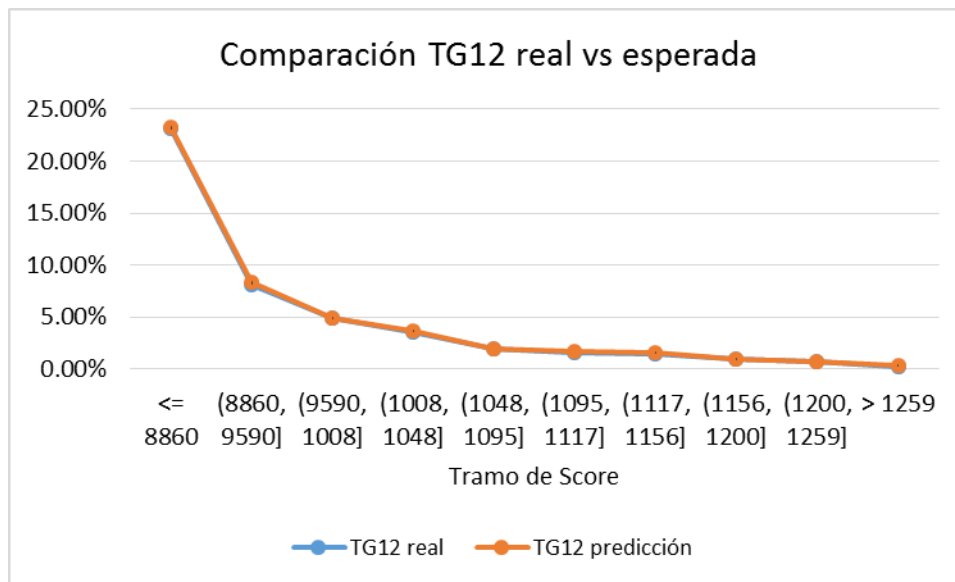


Imagen 40: Comparación gráfica de la TG_{12} por tramo de score (clientes sin abono de remuneraciones)

Y finalmente, el back test del modelo para este segmento se puede apreciar en la Imagen 41.

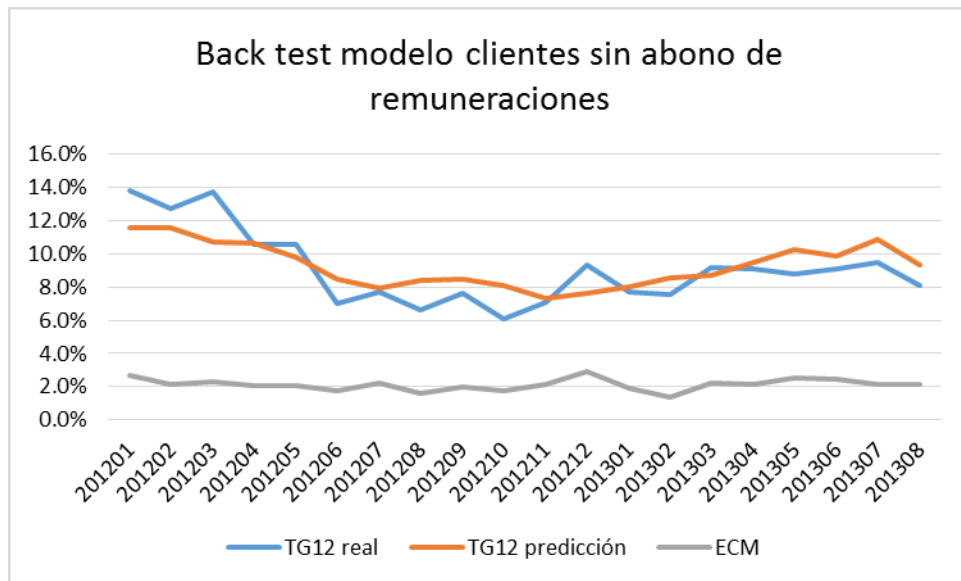


Imagen 41: Estabilidad del modelo a través del tiempo (clientes sin abono de remuneraciones)

4.8.3 Segmento 3: clientes con antigüedad ≥ 6 meses y con abono de remuneraciones en Bci (clientes sin abono de remuneraciones)

Para el segmento de clientes con abono de remuneraciones, la solución de la *scorecard* considera las siguientes variables:

Rango	β	Sig.	Score	Score máximo	Score mínimo
Edad <= 30	0.528	0.000	-65	0	-65
Edad > 30 and Edad <= 48	0.460	0.000	-57		
Edad > 48		0.000	0		
PIxPDI_New <= 0.020825	-1.914	0.000	236	236	0
PIxPDI_New > 0.020825 and PIxPDI_New <= 0.0507	-1.497	0.000	185		
PIxPDI_New > 0.0507 and PIxPDI_New <= 0.3327	-0.889	0.000	110		
PIxPDI_New > 0.3327 and PIxPDI_New <= 1.3333	-0.397	0.000	49		
PIxPDI_New > 1.3333		0.000	0		
R04 <= 153	2.823	0.000	-348	0	-348
R04 > 153 and R04 <= 174	1.983	0.000	-245		
R04 > 174 and R04 <= 229	1.044	0.000	-129		
R04 > 229 and R04 <= 244	0.483	0.000	-60		
R04 > 244		0.000	0		
ANTIG_CCT <= 12	0.692	0.000	-85	0	-85
ANTIG_CCT > 12 and ANTIG_CCT <= 24	0.667	0.000	-82		
ANTIG_CCT > 24 and ANTIG_CCT <= 36	0.382	0.000	-47		
ANTIG_CCT > 36		0.000	0		
Leverage > 0 and Leverage <= 4	-1.508	0.000	186	186	0
Leverage > 4 and Leverage <= 8	-1.182	0.000	146		
Leverage > 8 and Leverage <= 12	-0.692	0.000	85		
Leverage > 12 and Leverage <= 15	-0.359	0.000	44		
Leverage > 15		0.000	0		
Tiene_HIP_Sbif_pre_curse = 0	0.401	0.000	-49	0	-49
Tiene_HIP_Sbif_pre_curse = 1		0.000	0		
RentaUF <= 67	0.395	0.000	-49	0	-49
RentaUF > 67 and RentaUF <= 144	0.254	0.020	-31		
RentaUF > 144		0.000	0		
Constante	-4.422	0.000	1,186	1,186	1,186
				1,608	590

Tabla 24: Resultado de la regresión logística (clientes con abono de remuneraciones)

Este segmento muestra un *ECM* bastante bajo, entre el 0.06% y el 5.09%.

Tramo puntaje	Núm casos	Promedio ponderado TG_{12} real	Promedio ponderado TG_{12} predicción	Error predicción	Error cuadrático medio
$\leq 1,020$	11,672	9.93%	10.78%	-0.60%	5.09%
(1,020; 1,097]	9,251	2.98%	3.19%	0.13%	1.43%
(1,097; 1,153]	9,243	1.52%	1.58%	0.20%	0.70%
(1,153; 1,197]	8,457	1.37%	1.76%	-0.33%	0.66%
(1,197; 1,238]	9,495	0.84%	0.83%	0.11%	0.37%
(1,238; 1,284]	9,060	0.46%	0.40%	0.15%	0.20%
(1,284; 1,325]	10,067	0.34%	0.40%	0.07%	0.20%
(1,325; 1,371]	11,847	0.28%	0.23%	0.08%	0.13%
(1,371; 1,437]	12,902	0.14%	0.13%	0.07%	0.09%
> 1,437	14,184	0.06%	0.07%	0.04%	0.06%
Total	106,178	1.59%	1.71%	-0.01%	0.90%

Tabla 25: Predicción de la TG_{12} usando el score (clientes con abono de remuneraciones)

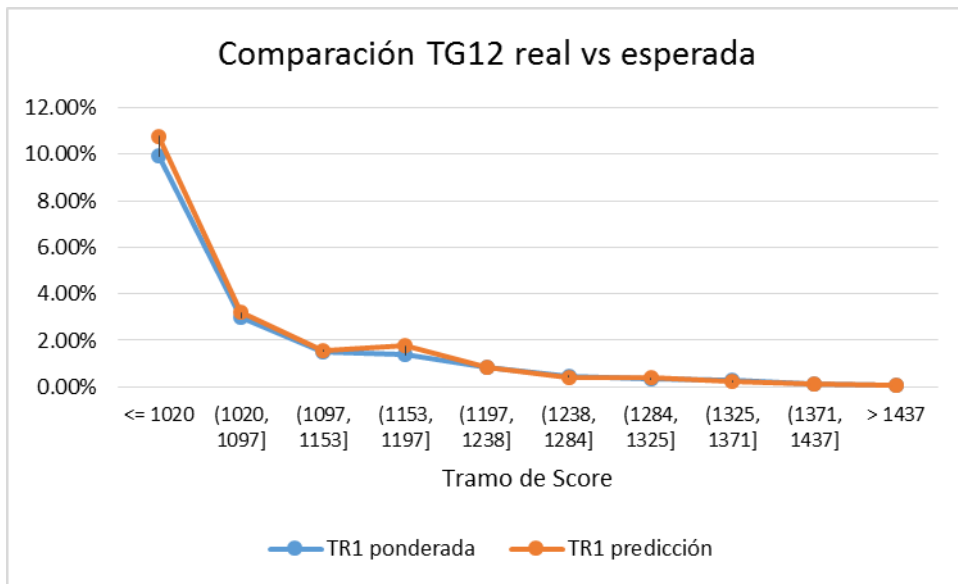


Imagen 42: Comparación gráfica de la TG₁₂ por tramo de score (clientes con abono de remuneraciones)

Finalmente, el modelo muestra estabilidad a lo largo del tiempo, tal como se muestra en la Imagen 43.

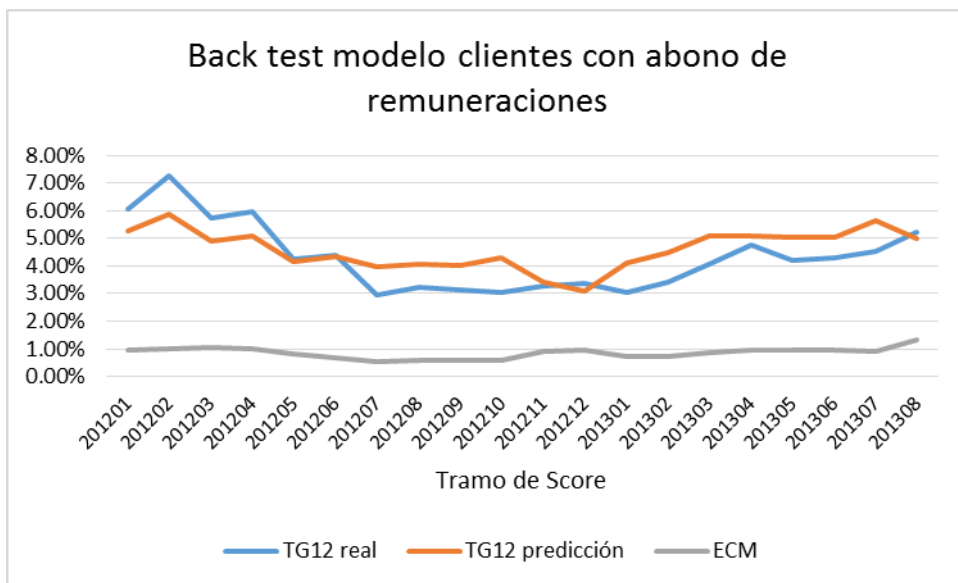


Imagen 43: Estabilidad del modelo a través del tiempo (clientes con abono de remuneraciones)

4.8.4 Segmento 4: colaboradores

El segmento de colaboradores es un caso especial, debido a que se cuenta con pocas observaciones y a que se cuenta con información restringida. Este segmento fue modelado utilizando la misma técnica de los segmentos anteriores, pero presentaba un $ECM > 13\%$. Se buscó entonces varias formas alternativas de modelar el problema en este segmento, con el objetivo de mejorar la precisión en la predicción.

El modelo definitivo fue el de un árbol de clasificación (con el algoritmo *exhaustive CHAID*)¹⁷, tal como se aprecia en la Imagen 44. Este modelo fue supervisado utilizando la variable TG_{12} (y con las mismas variables de entrada que el resto de los segmentos). El árbol mostraba un crecimiento en más niveles, pero se “podó” privilegiando un mejor soporte de cada nodo terminal.

Cada uno de los nodos terminales muestra una TG_{12} promedio, la cual se replica finalmente en cada uno de los colaboradores pertenecientes a dicho nodo. El ECM global del modelo es 0.95%, fluctuando entre un 0.057% y un 3.37% dependiendo del nodo.

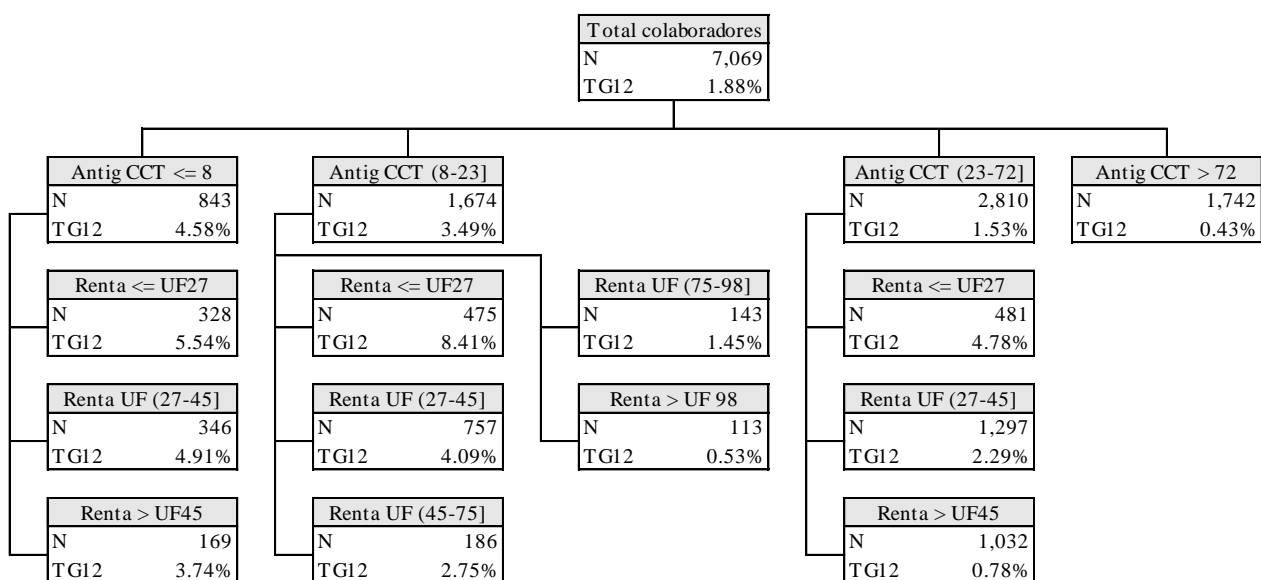


Imagen 44: Árbol de clasificación para el modelo colaboradores

Si bien se cuenta con historia suficiente, dado el bajo número de cursos por mes en este segmento (entre 300 y 500) el back test muestra distorsiones algo elevadas. Sin embargo, dada la magnitud

¹⁷ Se intentó modelar utilizando las siguientes técnicas: regresión logística (*scorecard*), regresión lineal, regresión no lineal generalizada, regresión lineal en logaritmos. El R^2 en estos modelos osciló entre el 10.3% y el 57%. Dada la baja bondad de ajuste, no se detallará la construcción de estos modelos.

del *ECM* se decide dar por validado el modelo, aun cuando, por razones administrativas, se decide también no implementar el modelo en este segmento.

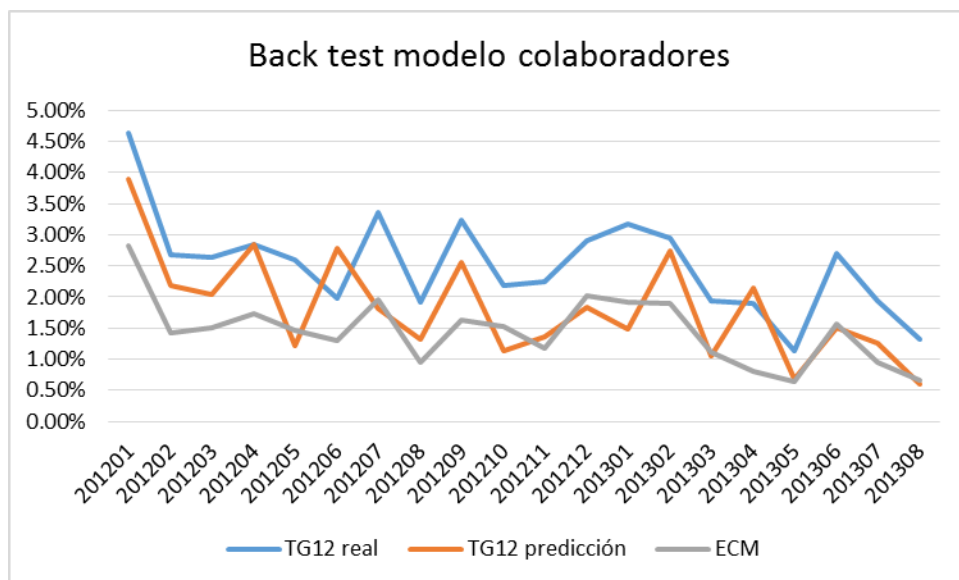


Imagen 45: Estabilidad del modelo a través del tiempo (colaboradores)

4.9 Aplicación de los modelos para establecer un esquema de precios mínimos

La Gerencia de Riesgo ha definido, en conjunto con la Gerencia de Pricing, un esquema de precios que soporten el nivel de gasto esperado en 12 meses. Se ha definido un *spread* mínimo a cobrar, diferenciado por segmento¹⁸.

La función para determinar el precio mínimo a cobrar se establece como:

$$Precio(X) \geq \overline{TG_{12}} + Spread(X)$$

Ecuación 10: Precio mínimo a cobrar por segmento

Donde *X* representa las características o atributos del cliente. Este esquema de precios se aplica directamente en la etapa de negociación del proceso de venta, permitiendo un margen de holgura en el cual el Banco puede negociar con sus clientes de acuerdo al riesgo que se espera y el gasto en riesgo que se espera asumir. Esto permite que cada curse nuevo mantenga un nivel de utilidad

¹⁸ La definición de segmento aquí difiere. Se trata del segmento comercial al cual pertenece cada cliente, determinado principalmente por el tipo de plan que tiene contratado el cliente. Por ejemplo, un cliente con un plan clásico pertenece al segmento “masivo”.

esperado mayor o igual a cero, y que a cada cliente se le cobre un precio diferenciado de acuerdo a su perfil de riesgo.

4.10 Aplicación de los modelos para establecer límites de endeudamiento óptimos

Finalmente, se consolidan los resultados de los 3 primeros segmentos (clientes nuevos, sin abono de remuneraciones y con abono de remuneraciones). Dado que la escala que se utilizó en la construcción de estos tres modelos es la misma, los *scores* obtenidos en los tres modelos son comparables entre sí.

Entonces se agrupan todos los clientes de estos tres modelos, y se determinan *clusters* en base al método K-Medias (cumpliendo con las características de homogeneizar a los miembros de un mismo *cluster*, pero diferenciarlos de los otros). Las variables de entrada para determinar los *clusters* que se consideran vienen dadas por una restricción del negocio: deben ser pocas (máximo dos) y deben ser de fácil entendimiento, replicabilidad y programación. Así, los segmentos obtenidos (7 en total) dependen del score obtenido en los modelos anteriores y la renta de los clientes¹⁹ (medida en Unidades de Fomento). En el Anexo 1 se muestra el resultado de la aplicación de la nueva segmentación para definir una estrategia *challenger* que puede ser aplicada a través de un piloto controlado. En esta tabla se muestra además un *leverage* máximo y una carga financiera (CF) máxima. El primero consiste en la cantidad de “veces renta” que el banco está dispuesto a endeudar a sus clientes (considerando la totalidad de deudas de consumo en el sistema financiero), y la carga financiera consiste en el porcentaje de la renta líquida mensual del cliente que destina para pagar obligaciones de crédito. La carga financiera se determina, de acuerdo a la actual política de créditos de Bci, como:

$$CF = \frac{2.5\% \times Deuda\ consumo + 5\% \times Deuda\ comercial + 0.7\% \times Deuda\ hipotecaria}{Renta\ líquida\ mensual}$$

Ecuación 11: Fórmula de cálculo de la carga financiera

Esta ecuación considera que, del total del saldo de la deuda de un cliente en el sistema financiero, cada categoría de deuda (consumo, comercial o hipotecaria) debe ser ponderada por un factor fijo que trata de simular el valor de las cuotas mensuales que el cliente cancela para cada categoría de deuda.

¹⁹ Si bien la renta puede parecer una variable redundante, el único modelo que utiliza esta variable es el segmento de clientes con abono de remuneraciones. Adicionalmente, para poder aplicar una estrategia *challenger*, se necesita utilizar esta variable, ya que la actual política *champion* la considera.

Por ejemplo, para cierto tramo de *score* y cierto tramo de renta, el nuevo modelo propondría endeudar a sus clientes hasta 8 veces renta, con un tope de 40% de carga financiera. Si en este mismo ejemplo, con una renta de \$400,000 mensual, una deuda de consumo de \$1,250,000 y una deuda hipotecaria de \$15,420,000, el actual *leverage* de este cliente sería de 3.125 veces y su carga financiera actual de 34.8%. De acuerdo al modelo propuesto, el cliente podría llegar a endeudarse hasta \$3,200,000 en consumo. Como ya tiene una deuda actual de consumo por \$1,250,000, su capacidad incremental de deuda (que se traduce en una eventual oferta) es de \$1,950,000. Sin embargo, el tope de carga financiera limita esta exposición. El cliente cuenta con una carga financiera disponible de 5.2% (esto es, el tope de 40% menos la actual CF de 34.8%), representando un valor monetario de \$20,810 (5.2% multiplicado por la renta de \$400,000). Si asumimos una tasa de interés del 1.3% y un plazo de 60 meses, al calcular el valor presente de un pago equivalente a estos \$20,810 se obtendría un posible crédito de hasta \$863,269. Entonces, a este cliente se le calcularía una oferta por \$863,269 a través de esta nueva propuesta.

Ahora bien, esta propuesta fue diseñada considerando que el nivel de TG_{12} esperado con la propuesta deba ser similar al de la situación actual. No se busca disminuir el nivel de gasto en riesgo, sino que, con el mismo nivel que se tiene actualmente, ser más preciso en la asignación de límites de endeudamiento. El algoritmo para determinar el nivel óptimo de *leverage* y carga financiera máximos es el siguiente:

- Considerar la base de desarrollo de los modelos anteriormente descritos.
- Aperturar cada *cluster* por tramos de *leverage* alcanzado posterior al *course*, considerando tramos marginales (de 1 en 1).
- Acumular el gasto en riesgo (como porcentaje) de cada uno de los tramos de *leverage*, para cada uno de los *clusters*.
- Considerar el *leverage* máximo permitido de manera de igualar la misma TG_{12} total del *cluster* o al menos encontrar un valor similar.
- Repetir el mismo procedimiento, pero aperturando por tramos marginales de carga financiera (de 5% en 5%).

Por ejemplo, para el *cluster* de tramo de *score* $\geq 1,150$ y tramo de renta entre UF 21 y 49, se tiene que²⁰:

	Tramo de <i>leverage</i>								Total
	<2	2-4	4-6	6-8	8-10	10-12	12-14	14-16	
TG_{12} acumulada	1.64%	1.13%	1.29%	4.53%	4.28%	4.50%	4.41%	4.41%	4.50%

Tabla 26: Ejemplo de determinación del *leverage* óptimo

²⁰ En este ejemplo se tramificó el *leverage* de 2 en 2 sólo para ajustar la presentación de la tabla en la página y para hacer más didáctica la explicación.

Como el nivel de TG_{12} del *cluster* total es 4.5%, lo que se desea es encontrar el nivel de *leverage* que se aproxime a este total. En este caso, cuando se endeuda a los clientes hasta 12 veces renta se alcanza el mismo nivel de TG_{12} , siendo este el *leverage* máximo permitido para éste.

El ejercicio se replica para la carga financiera de la misma forma, en los 21 *clusters*.

4.11 Piloto

Ya teniendo en clara la definición de la política de endeudamiento y los precios mínimos a cobrar, lo que queda es diseñar un piloto que pueda poner a prueba esta nueva metodología a través de una prueba de *champion-challenger*.

4.11.1 Diseño del piloto

Mensualmente, el proceso de generación de campañas masivas evalúa el total de la cartera para determinar ofertas de campaña. El primer paso, luego de consolidada la base inicial con la que se calcularán las ofertas, es filtrar la base. Para ello, se aplican los denominados “filtros de riesgo”, que equivalen a ciertos atributos que los clientes tienen y que impiden que puedan operar con Bci. Por ejemplo, un cliente que actualmente posee moras, o que ha realizado fraude no es sujeto de crédito para Bci. Adicionalmente, existen otros filtros denominados “filtros de perfil”, que consisten en atributos del perfil objetivo de Bci. Si bien un cliente que tiene una renta líquida mensual inferior a \$400,000 no implica un riesgo de operar, no forma parte del mercado objetivo de Bci. Por confidencialidad, no se mostrará el listado completo de filtros de riesgo ni de perfil que Bci aplica en la generación de las campañas masivas.

Una vez seleccionados los clientes que pertenecen al mercado objetivo y que pasan los filtros de riesgo, el siguiente paso es seleccionar aleatoriamente un 10% de los potenciales clientes a ofertar como grupo de tratamiento (o *challenger*), siendo el 90% restante el grupo de control (o *champion*). Es importante mencionar que esta selección se realiza una sola vez, al inicio del piloto. Es decir, un cliente que en $t = 1$ fue asignado al grupo de tratamiento no puede aparecer en periodos posteriores en el grupo de control y viceversa. Naturalmente, de todas maneras, como existen clientes que dejan de serlo y clientes que se vienen incorporando en cada periodo, el grupo de tratamiento va a ir disminuyendo como proporción respecto del total. Para mantener esta proporción constante, se selecciona aleatoriamente un número de clientes para completar el 10% de entre los nuevos clientes que van apareciendo mes a mes (es decir, que no estaban asignados ni al grupo de tratamiento ni al grupo de control). Finalmente, sólo resta calcular el nivel de endeudamiento máximo. Para el grupo de tratamiento, la metodología aplicada es la propuesta en el Anexo 1, mientras que para el grupo de control se aplica la metodología ya existente.

4.11.2 Criterios de éxito del piloto

Paralelamente se definen también los criterios de éxito del piloto. Para ello, se ha acordado en conjunto con los distintos gerentes y sub gerentes los siguientes KPI:

- El *Win Rate*²¹ del grupo de tratamiento debe ser igual o mayor al del grupo de control. Si no se cumple esta condición, el piloto se clasifica como no exitoso.
- La TG_{12} del grupo de tratamiento debe ser similar a la del grupo de control. Si este indicador es menor en el grupo de tratamiento, significa que existe espacio para poder colocar más ofertas, haciendo que el piloto sea insuficiente. Por el contrario, si el indicador es menor en el grupo de control, significa que el piloto es más riesgoso de lo esperado, determinando el fracaso del mismo.
- Dado que la TG_{12} se obtendrá recién al doceavo mes de madurez de las camadas, se seguirá también un indicador de mora. Este indicador de mora marca a las operaciones que tienen al menos 30 días de mora, y se siguen con la misma lógica de madurez de camadas. Un buen punto de referencia es utilizar el indicador de mora al cuarto mes de madurez de las camadas. Para que el piloto sea exitoso, se espera que el indicador de mora al cuarto mes del grupo de tratamiento sea menor o igual que en el grupo de control.

4.11.3 Aprobación y ejecución del piloto

Este piloto así diseñado fue presentado a un comité directivo del banco, compuesto por la alta gerencia de Bci. Este comité decidió autorizar el piloto y ponerlo en ejecución desde marzo de 2015. Semanalmente se entregan informes de avance del piloto para medir los KPI indicados anteriormente. A la fecha de elaboración de este documento el piloto ha mostrado indicadores de desempeño satisfactorios y sobresalientes, lo que ha llevado a plantear un plan para realizar el *roll-out* del piloto en el 100% de la cartera.

4.12 Generalización

Si pensamos en un problema de esta índole, lo que se busca es predecir cierto nivel de costos, para de esa forma fijar precios mínimos y asegurar una oferta adecuada. El dominio de este problema abarca entonces cualquier problema donde lo que se pretenda es, a través de un control adecuado de los costos, establecer modelos de venta proactivos de servicios que integren tanto la cantidad del servicio otorgable como el precio a cobrar por él. En la Imagen 46 se puede apreciar el diagrama del dominio del problema de generalización, donde se parte del problema

²¹ El *Win Rate* (*WR*) consiste en la proporción de la oferta que finalmente se vende. Se calcula como $WR = \frac{\text{Total venta}}{\text{Total oferta}}$ y se mide como porcentaje.

más específico (predicción del gasto en riesgo en ofertas proactivas de créditos de consumo) hasta el problema más general detectado, que consiste en la predicción de costos.

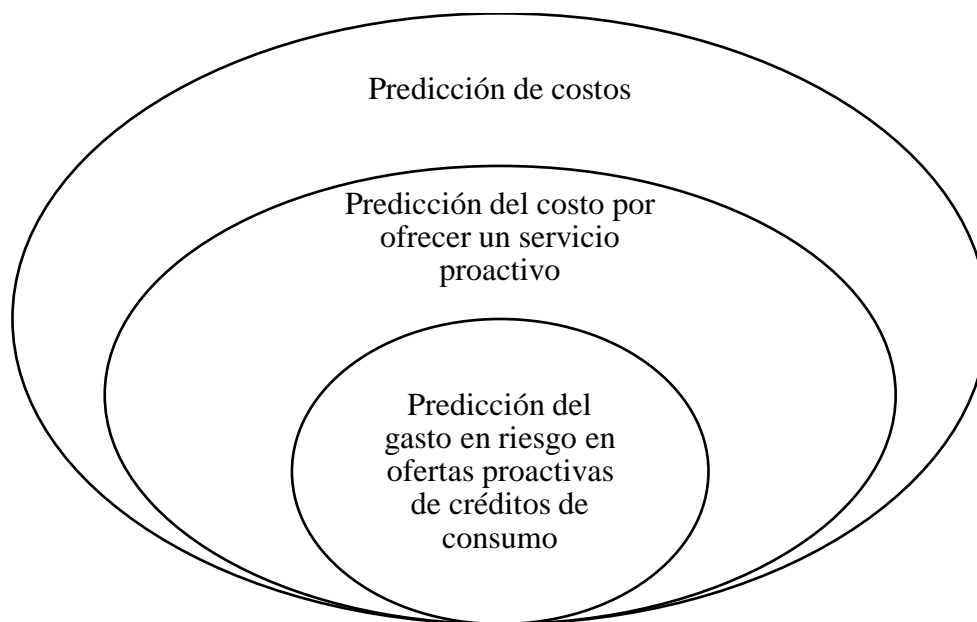


Imagen 46: Diagrama del dominio del problema

La predicción de costos por sí sola no sirve de mucho a una empresa, sino hasta que se utiliza para dirigir las estrategias de fijación de precios mínimos por sobre el costo y la determinación de la cantidad óptima a ofrecer del servicio. Algunos usos potenciales que se le pueden dar a una solución de esta índole (ver Imagen 47) vienen dados por los siguientes ejemplos:

- En la industria de salud (clínicas privadas) un modelo de este tipo puede servir para predecir el costo de un seguro de salud (medido como el costo de utilizar los servicios clínicos), que sirva para determinar el precio mínimo a cobrar a los pacientes y la cobertura óptima.
- En la industria del *retail* se puede predecir el costo de que un cliente deje de pagar su tarjeta de crédito, para así fijar comisiones mínimas e intereses moratorios, así como también determinar el cupo a ofrecer en dichos productos.
- En la industria de seguros se puede predecir el costo de pagar las pólizas de seguros para así fijar precios mínimos a cobrar a los asegurados, así como también la cobertura óptima de dichas pólizas.
- En telecomunicaciones se pueden establecer tarifas mínimas de acuerdo a la predicción del costo de no pago de un cliente, así como también la cantidad óptima de servicios a ofrecer (minutos, por ejemplo).

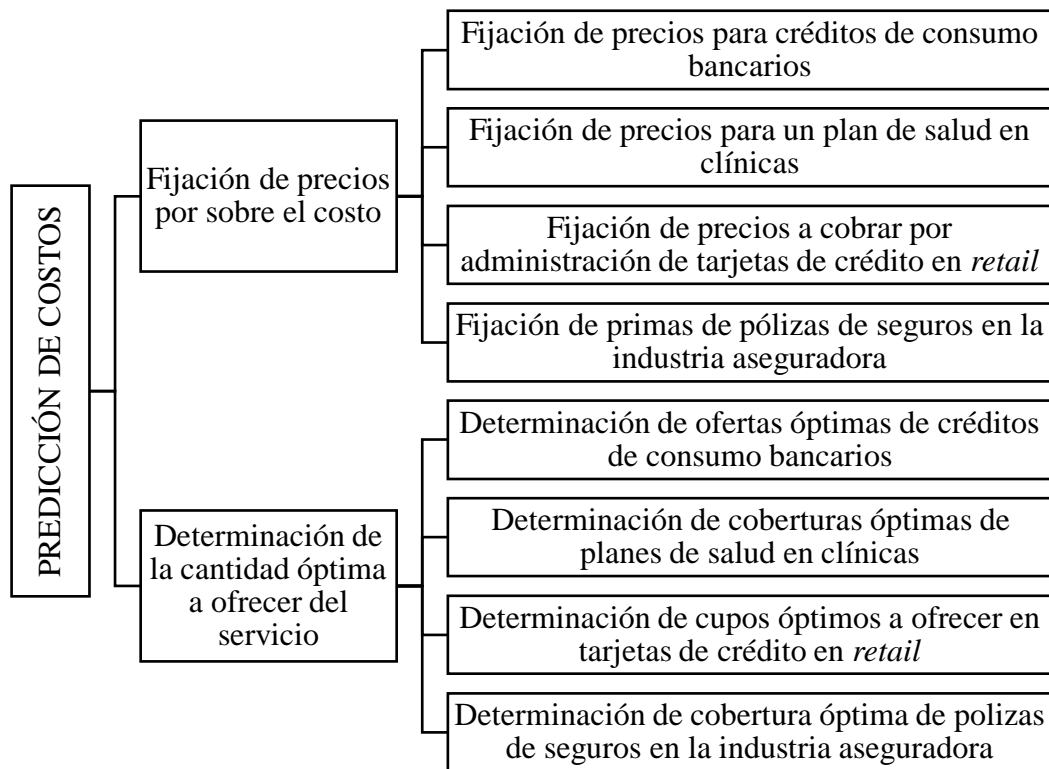


Imagen 47: Ejemplos de uso de una solución generalizada

Como se aprecia, la generalización de este problema es factible y replicable en otras industrias, haciendo que el proyecto pueda trascender más allá del ámbito aquí expuesto.

5 EL PROTOTIPO

El proyecto de tesis tiene asociado el desarrollo e implementación de un sistema tecnológico de apoyo que dé sustento o facilite todo el trabajo que hay por detrás. El sistema que se construyó, a modo de prototipo, realiza las acciones anteriormente descritas en el diagrama de pistas para Ejecución de modelos de venta (ver título 3.4.2). Se representará a continuación el prototipo que realice las tareas descritas en el diagrama de pistas mediante notación UML.

5.1 Diagramas UML

5.1.1 Diagrama de paquetes

El diagrama de pistas para Ejecución de modelos de venta reside en un sistema que contendrá un paquete de interfaz, que permita relacionar al mismo con el usuario a través de una plataforma web. Por simplicidad no se mostrarán las actividades relacionadas a la administración del sistema (por ejemplo, relacionadas con la carga de los modelos) ni aquellas relacionadas con el control de usuarios.

Este sistema utilizará el paquete `ControlAccesoDatos`, el cual contiene la clase `ejecutaModeloDAO`, que contiene las reglas de negocio asociadas a la consulta de modelos disponibles.

También se utilizará el paquete `AccesoDatos`, que contiene las clases `ejecutaModeloDTO`, encargada de procesar la petición de acceso a los datos de los modelos almacenados en la base de datos y `datosModeloDTO` que contiene los datos asociados a los clientes sobre los cuales se desea aplicar los modelos de minería de datos.

La Imagen 48 muestra el diagrama de paquetes del sistema. Todos estos paquetes fueron utilizados en la construcción del prototipo del sistema.

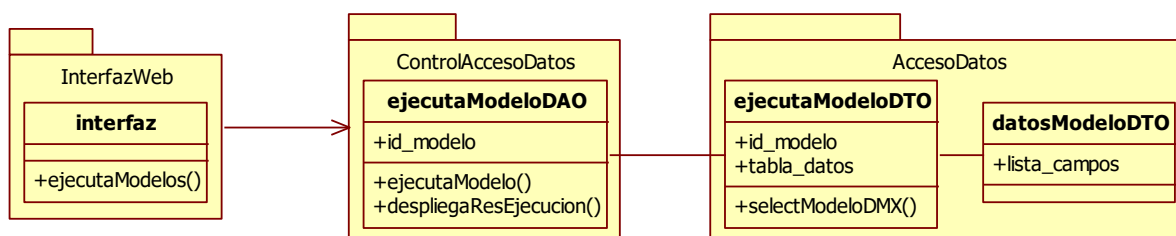


Imagen 48: Diagrama de paquetes del sistema

5.1.2 Diagrama de casos de uso

El único participante que interactúa con el sistema en la actividad mencionada es el Analista de estrategias, que puede solicitar la ejecución de los modelos. Por otra parte, el sistema puede Ejecutar modelos de forma autónoma o programada, de acuerdo a las necesidades de lanzamiento de campañas periódicas. Estos casos de uso se pueden ver en la Imagen 49.

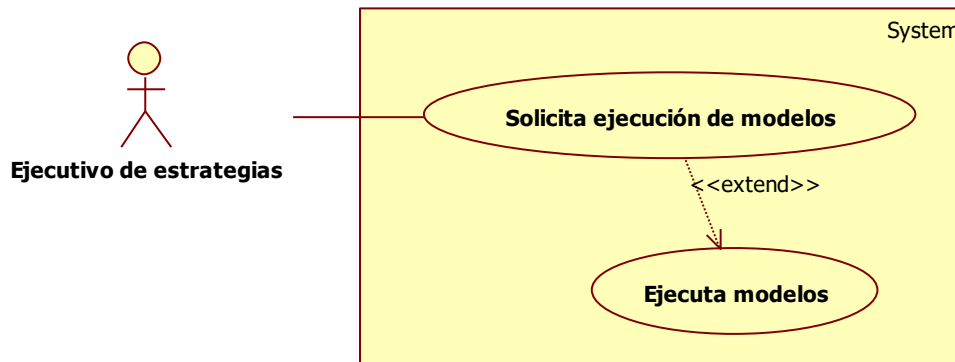


Imagen 49: Diagrama de casos de uso del sistema

5.1.3 Diagramas de secuencia

Por motivos de simplicidad, se mostrará el diagrama de secuencia asociado a solicita ejecución de modelos, el cual extiende la actividad Ejecuta modelos.

El Ejecutivo de estrategias (usuario) invoca la ejecución de los modelos a través de la Interfaz. Ésta envía la instrucción al sistema. El sistema toma esta solicitud y revisa todos los modelos vigentes que estén almacenados en su base de datos. Luego, para cada modelo solicita su ejecución al motor de base de datos. En cada una de estas iteraciones el sistema va guardando el resumen de la ejecución, para luego informar todos estos resúmenes juntos a la interfaz. Este diagrama se puede apreciar en la Imagen 50.

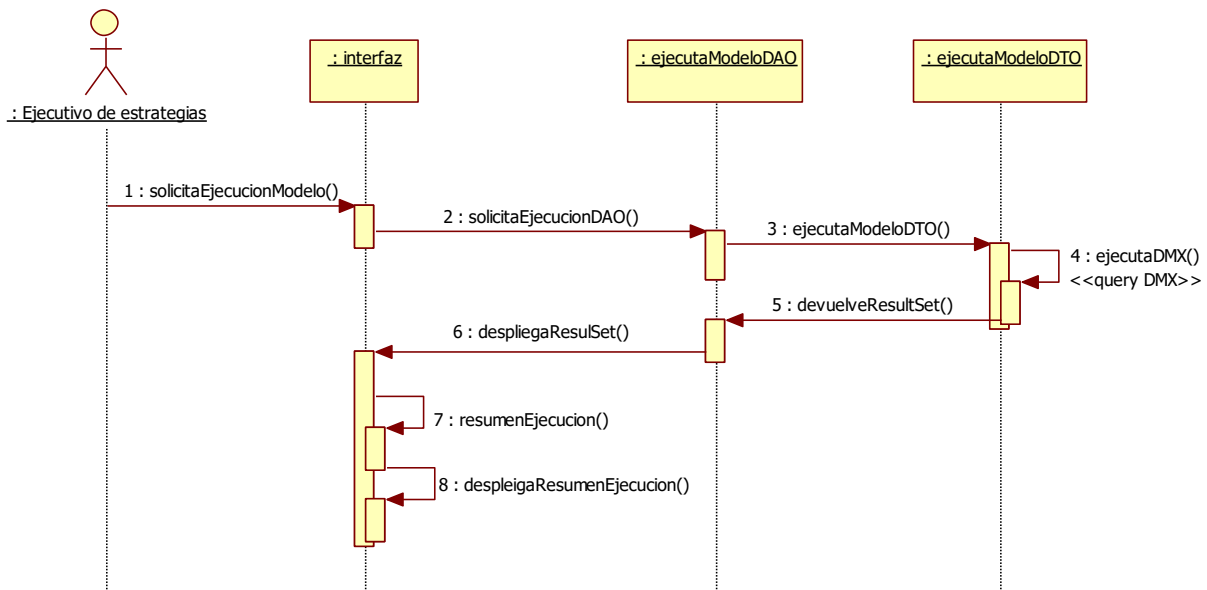


Imagen 50: Diagrama de secuencia para Solicita ejecución de modelos

5.2 Tecnologías de apoyo

Ya establecido el ámbito y los sistemas involucrados en el proyecto, es necesario identificar las tecnologías que soportan la ejecución de estos procesos. A continuación se enumeran algunas de las tecnologías existentes que permiten la construcción del prototipo tecnológico.

- **Microsoft SQL Server**

Microsoft SQL Server es un sistema para la gestión de bases de datos producido por Microsoft basado en el modelo relacional. Esta tecnología es la que, en conjunto con Teradata, se utiliza en la actualidad para manejar las bases de datos del Banco. La ventaja principal es que la administración es descentralizada, es decir, se cuenta con diversos servidores con permisos diferenciados (escritura, lectura de datos) administrable por los mismos usuarios. Por lo mismo, la principal desventaja radica en la dependencia de la *performance* física de la máquina donde está alojado el servidor de datos.

- **Teradata Database**

Esta tecnología es un sistema para la administración del *Data Ware House* (DWH) corporativo comercializado por Teradata Corporation. Es administrada centralmente por una unidad de sistemas de Bci y su principal característica diferenciadora es la capacidad para almacenar y ejecutar procesos para enormes cantidades de datos debido al paralelismo en la ejecución de consultas de datos. Se basa en el modelo relacional y actualmente es la fuente oficial de datos de Bci. De aquí se alimentan distintos *Data Marts* (DM). La principal ventaja radica en la rapidez, pero cuenta con la gran desventaja

de que cualquier cambio (por ejemplo, creación de un modelo de datos relacional para alojar el prototipo) debe pasar por las áreas que administran esta herramienta, haciendo poco administrable un eventual prototipo.

- **ASP.NET**

ASP.NET es un *framework* para aplicaciones web desarrollado y comercializado por Microsoft. Es usado por programadores para construir sitios web dinámicos, aplicaciones web y servicios web XML. Apareció en enero de 2002 con la versión 1.0 del .NET Framework, y es la tecnología sucesora de la tecnología Active Server Pages (ASP). ASP.NET está construido sobre el Common Language Runtime, permitiendo a los programadores escribir código ASP.NET usando cualquier lenguaje admitido por el .NET Framework (por ejemplo, VB.NET, C++ ó C#). Esta tecnología permite construir la interfaz de usuario que conectará el sistema con los datos residentes en bases de datos que se administran en Bci.

- **Microsoft Analysis Services**

Microsoft Analysis Services (MSAS) es la solución de Microsoft para la construcción de modelos de minería de datos, y tiene la ventaja de que se integra casi de manera nativa con Microsoft SQL Server, ejecutando modelos de minería de datos a través de la misma interfaz y con la posibilidad de automatizar su ejecución.

- **IBM SPSS Statistics y SPSS Modeler**

SPSS Statistics es un software estadístico comercializado por IBM que se caracteriza por su interfaz simple, amigable y limpia. Permite el desarrollo de modelos de minería de datos de inicio a fin. SPSS Modeler es el sucesor de Clementine, también desarrollado por IBM, que permite ejecutar procesos de una forma visual y automática.

- **RapidMiner**

Es un software opensource dedicado a la construcción y evaluación de modelos de minería de datos. También permite la ejecución de procesos de análisis de datos.

5.3 Arquitectura tecnológica

De cara a la implementación del prototipo tecnológico, se sugiere aprovechar las tecnologías existentes en Bci para habilitar los procesos que se requiere. Por esta razón, es que se presenta el siguiente diagrama de arquitectura (Imagen 51), en el que se muestra un modelo de arquitectura de 3 capas (cliente, servidor de aplicaciones y servidor de datos) para un sistema corriendo sobre una plataforma web. Referente a las tecnologías específicas que se utilizará en cada capa, como sistemas operativos, motores de bases de datos o lenguaje de programación, se

sugieren sólo a modo de ejemplo, y es una buena aproximación de lo que se podría implementar como solución final.

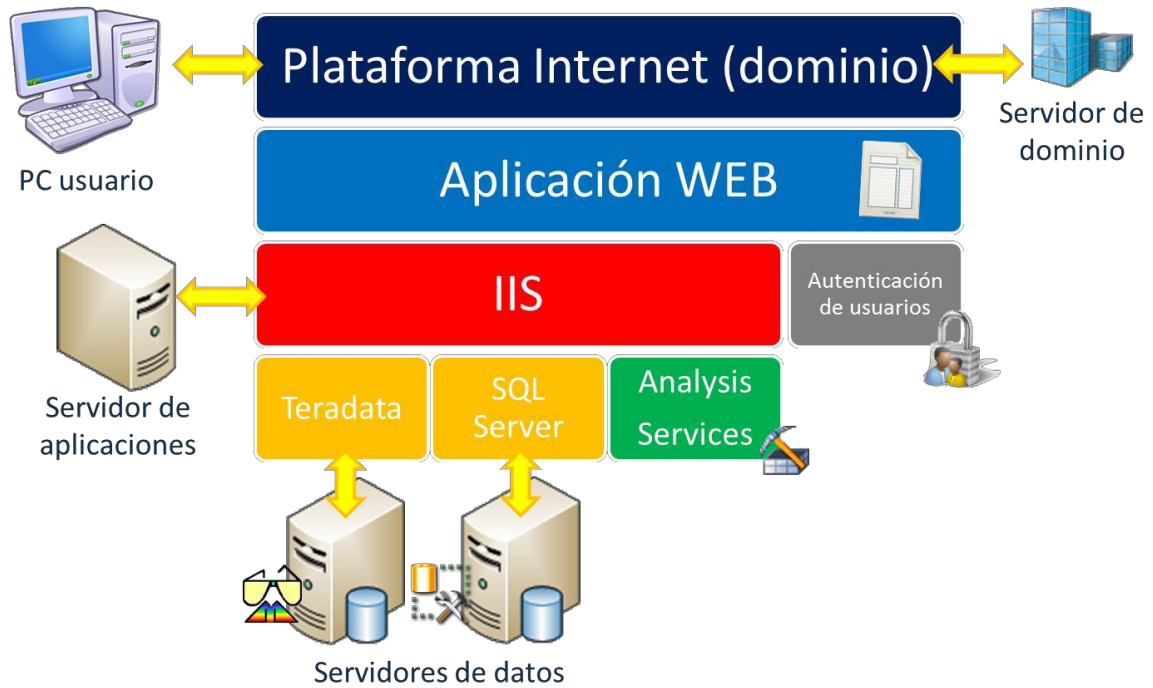


Imagen 51: Diagrama de arquitectura tecnológica sugerida para el prototipo

6 CONCLUSIONES

La competencia por una mayor utilidad en la industria bancaria históricamente viene dada por el volumen de ventas. En un mercado relativamente pequeño como el chileno ya no hay mucho más espacio para crecer a través de esta vía. Para diferenciarse de la competencia, Bci ha impulsado un Plan Corporativo de Transformación de Riesgo que cambia el paradigma de maximizar la utilidad hacia una cultura de reducción de costos, principalmente el gasto en riesgo.

Existen cuatro puntos principales que se abordaron con el desarrollo de este proyecto: rediseño del proceso de generación de campañas, establecimiento de un modelo de predicción del gasto en riesgo para la venta de créditos de consumo, una nueva metodología de asignación de precios ajustados por riesgo y optimización del límite de endeudamiento de los clientes. Referente al primer objetivo se han impulsado diversas mejoras que han llevado a disminuir los tiempos de respuesta, automatizar la generación de ofertas, disminuir el error humano y prácticamente eliminar el tiempo de distintos actores destinado a tareas que no aportan valor al negocio. Un mejor modelo de predicción del gasto en riesgo para la venta permite una mejor asignación de recursos al redistribuir los costos, liberar dinero inmovilizado por constitución de provisiones y anticipar posibles pérdidas por castigo de créditos. La situación sin proyecto implicaba un cobro indiferenciado de precios a los clientes, obteniendo utilidad en algunos y pérdidas en otros, mientras que la situación con proyecto corrige los precios y logra que a cada cliente se le cobre un precio diferenciado de acuerdo a su perfil de riesgo. Finalmente un esquema de límites de endeudamiento óptimos permite no sobre endeudar a los clientes, facilitando un mejor comportamiento de pago y posiblemente mejorando el riesgo de la cartera.

Las técnicas y métodos utilizados en este proyecto son consistentes con aquellas aprendidas en el MBE. Se trató de abarcar el proyecto de forma integral viéndolo desde distintas perspectivas: procesos, personas, analítica, conocimiento del negocio y financiero, de acuerdo al contexto actual de la organización.

Los indicadores de desempeño de los modelos llevados a cabo gracias a este proyecto, si bien son anticipados aún, auguran un éxito de mediano-largo plazo del piloto aplicado, éxito que puede ser replicable al resto de la cartera. Adicionalmente, un posible esquema de *framework* permitiría generalizar este tipo de proyecto en otro tipo de empresas con problemas similares, apoyando el desarrollo del trabajo en una herramienta tecnológica que facilite su implementación.

Luego de concluido el proyecto, las enseñanzas principales que se rescatan tienen más relación con el factor humano, clave para una comunicación y persuasión efectiva. El proyecto también le ha permitido al autor desarrollar un proceso de cambio personal y laboral que le han abierto caminos que antes no tenía a su disposición.

Como observación final, es deber del autor señalar que este proyecto y su autor han sido premiados al interior de la corporación Bci, ya que le ha significado a la organización aterrizar un nuevo paradigma que la alta gerencia quería materializar en un corto tiempo de trabajo y con un alto estándar de calidad y aceptación.

7 BIBLIOGRAFÍA

- Al Zabir, O. (2008). *Building a Web 2.0 Portal with ASP.NET 3.5*. O'Reilly.
- Anderson, J. (2013). *The Lean Change Method: Managing Agile Organizational Transformation Using Kanban, Kotter and Lean Startup Thinking*. CreateSpace Independent Publishing Platform.
- Baesens, B., Setiono, R., Mues, C., & Vanthienen, J. (2003). Using neural network rule extraction and decision tables for credit-risk evaluation. *INFORMS Management Science Review*, Vol. 49(3), 312-329.
- Barros, O. (2012). *Ingeniería de Negocios*. Departamento Ingeniería Industrial, Universidad de Chile. Obtenido de <http://blog.obarros.cl>
- Bci. (2012). *Memoria Anual*. Recuperado el 2014, de <http://www.bci.cl/medios/2012/investor/dectos/memorias/BciMemoria2012.pdf>.
- Bci. (2014). *Memoria Anual*. Recuperado el 2014, de http://www.bci.cl/medios/2012/investor/dectos/memorias/Memoria_Bci_2014.pdf.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, P. J. (1984). *Classification and Regression Trees*. Wadsworth International Group.
- Codd, E. (1970). *A Relational Model of Data for Large Shared Databanks*. Communications of the ACM.
- Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines*. Cambridge University Press.
- Defu, Z., Zhou, X., Leung, S., & Zheng, J. (2010). Vertical bagging decision trees models for credit scoring. *Journal of Expert Systems With Applications*, Vol. 37, 7838-7843.
- González, M., & León, J. (2007). Análisis del endeudamiento de los hogares colombianos. *Revista Desarrollo y Sociedad*(60), 127-154.
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., . . . Sriram, S. (2006). Modelling Customer Lifetime Value. (U. o. Maryland, Ed.) *Journal of Service Research*, Vol. 9(2), 139-155.

- Hax, A. (2010). *The Delta Model*. Springer.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2da. edición ed.). Wiley.
- Johnson, M., Christensen, C., & Kageman, H. (2008). Reinventing your Business Model. *Harvard Business Review*.
- Kaplan, R., & Norton, D. (1996). *The Balanced Scorecard: Translating Strategy Into Action*. Harvard Business School Press.
- Khalesi, N., & Shokuhi, A. H. (2011). Evaluation of data mining classification techniques and performances for banking customers credit scoring. *5th symposium on advances in science and technology (SASTech)*. Khavaran Higher-education Institute.
- Lai, T. L., Robbins, H., & Wein, C. Z. (1978). Strong consistency of least squares estimates in multiple regression. *PNAS*, Vol. 75(7), 3034-3036.
- MacLenna, J., Tang, Z., & Crivat, B. (2009). *Data Mining With Microsoft SQL Server 2008*. Wiley Publishing.
- Meera, R., & Tulasi, B. (2015). Credit Scoring Process using Bankinkg Detailed Data Store. *International Journal of Applied Information Systems (IJAIS)*, Vol. 8(6), 13-18.
- Ngai, E., Xiu, L., & Chau, D. (2009). Application of data mining techniques in customer relationship management: a literature review and classification. *Journal of Expert systems With Applications*, Vol. 36, 2592-2602.
- Ostelwalder, A., & Pigneur, Y. (2010). *Business Model Generation*. Wiley & Sons.
- Porter, M. (1996). What is strategy. *Harvard Business Review*.
- Ruiz-Tagle, J., García, L., & Miranda, Á. (2013). Proceso de endeudamiento y sobre endeudamiento de los hogares en Chile. *Documentos de trabajo N° 703*.
- Sapag, N. (2007). *Proyectos de inversión: formulación y evaluación* (1a edición ed.). Pearson Education de México.
- Sapag, N., & Sapag, R. (2008). *Preparación y evaluación de proyectos* (5a edición ed.). McGraw-Hill.

Schebesch, K., & Stecking, R. (2005). Support Vector Machines for Classifying and Describing Credit Applicants: Detecting Typical and Critical Regions. *Journal of Operations Research Society*, Vol. 56(8), 1082-1094.

Schmuller, J. (2000). *Aprendiendo UML en 24 horas*. Prentice Hall.

Torres Moncayo, J. (2009). *Lean production: cómo llegar a ser lean sin mucho esfuerzo*. ITESM.

White, S., & Miers, D. (2009). *Guía de referencia y modelado BPMN. Comprendiendo y utilizando BPMN*. Future Strategies.

Zhang, H. (2004). *The optimality of Naïve Bayes*. Proc. FLAIRS.

8 ANEXOS

8.1 Anexo 1: tabla de endeudamiento propuesta según modelos (segmentos 1 al 3)

Tramo de renta	Tramo de score													
	< 881		882 – 930		931 – 954		955 – 978		979 – 1,086		1,087 – 1,149		≥ 1,150	
	Lev máx	CF máx	Lev máx	CF máx	Lev máx	CF máx	Lev máx	CF máx	Lev máx	CF máx	Lev máx	CF máx	Lev máx	CF máx
< UF 21	0	0%	4	30%	6	40%	8	40%	8	40%	8	40%	8	40%
UF 21-49	0	0%	4	30%	6	40%	10	45%	10	45%	11	50%	12	60%
≥ UF 49	0	0%	4	30%	6	40%	12	60%	14	60%	16	60%	18	60%