



**UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERIA INDUSTRIAL**

**ENRIQUECIMIENTO DE UN MODELO CREDITICIO A CONSUMIDORES CON  
VARIABLES TRANSACCIONALES**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL**

**ARIEL RAMOS GUTIÉRREZ**

**PROFESOR GUÍA:  
JOSÉ MIGUEL CRUZ GONZÁLEZ**

**MIEMBROS DE LA COMISIÓN:  
CRISTIÁN BRAVO ROMÁN  
RICHARD WEBER HASS**

**SANTIAGO DE CHILE  
OCTUBRE 2012**

RESUMEN DE LA MEMORIA PARA OPTAR AL  
TITULO DE INGENIERO CIVIL INDUSTRIAL  
POR: ARIEL RAMOS GUTIÉRREZ  
FECHA: 30/10/2012  
PROF. GUIA: SR. JOSÉ MIGUEL CRUZ G.

## **ENRIQUECIMIENTO DE UN MODELO CREDITICIO CON VARIABLES TRANSACCIONALES.**

El trabajo de esta memoria consistió en analizar la información transaccional de una empresa del *retail* para la generación de variables de carácter transaccional, que permitan entender si existe un impacto de este tipo de variables cuando son incorporadas en un modelo de riesgo crediticio. Estos tipos de modelos son utilizados en la industria para gestionar carteras de clientes de acuerdo a un conjunto de características, buscando identificar a los grupos más riesgosos de clientes de acuerdo a características particulares observadas.

El acceso al crédito por parte de nuevos sectores de la población, así como el constante cambio en los patrones de comportamiento de los clientes, obligan a la industria a estar constantemente perfeccionando sus modelos de gestión de riesgo a fin de poder capturar a tiempo cualquier cambio en los patrones de consumo de los clientes. Es por este motivo, que este trabajo se orienta a la exploración de un nuevo conjunto de variables y analiza como éstas impactan en el actual modelo utilizado por la empresa.

La metodología empleada para realizar este trabajo utiliza técnicas provenientes de la minería de datos para la extracción, transformación y construcción de las variables relevantes para el modelo. Por otro lado, la fase de modelamiento de las variables seleccionadas utiliza como técnica la regresión logística, esta herramienta permite el modelamiento de variables dependientes con características dicotómicas y es la herramienta *ad-hoc* dada la naturaleza del problema abordado.

El trabajo propone un modelo mixto, que incorpora parte de las variables de comportamiento ya utilizadas por la empresa y lo complementa con un conjunto de cuatro variables de carácter transaccional identificadas para este trabajo.

Los resultados obtenidos con los distintos modelos desarrollados arrojaron que las variables transaccionales tienen impacto en el modelo crediticio y ayudan a mejorar el poder de discriminación del modelo en uso por la empresa, complementando el actual modelo. Por otro lado, las variables seleccionadas proveen a la empresa de una nueva fuente de información donde buscar elementos que ayuden a perfeccionar su modelo crediticio.

## **Agradecimientos**

Al escribir estas palabras, se vienen a mi cabeza las innumerables veces que imaginé el día en que llegaría este momento y cuáles debían ser las palabras precisas que lo describieran.

La verdad, es que desde un comienzo, el paso por la universidad se presentó como un constante desafío, donde a pesar de haber tropezado en muchas ocasiones, siempre tuve a un amigo al lado en el cuál apoyarme y poder superar el escollo que se presentaba. En este grupo quiero destacar a mis amigos del colegio Osmar, César Johan y Sebastián, con quienes he compartido por más de doce años y con los cuales he podido forjar una amistad que ha soportado los años y las distancias. También quiero mencionar a los amigos que descubrí en mi paso por la universidad: Ignacio, Anita, “la Miranda”, Gonzalo y especialmente Álvaro con quien me tocó compartir en largas jornadas de estudio y apoyarnos durante largo tiempo de nuestra vida universitaria. Mención especial para Valeria, con quien compartí en los inicios de la vida universitarios y fue fundamental para superar los duros escollos de los primeros años universitarios. Finalmente quiero agradecer a mis amigos del CAIN Mathias y Pablo, quienes a pesar de lo disímiles de nuestros intereses, siempre han sabido darse un tiempo para compartir una conversación reconfortante.

También debo agradecer a mis profesores de la memoria, José Miguel y Cristián ya que sin su ayuda y consejo no hubiera podido sacar esta tarea adelante. Al equipo de Ripley por su buena onda y la buena acogida que me dieron para realizar este trabajo, especialmente a Erika y Claudio, ya que sin su apoyo no hubiera podido sacar esta memoria adelante.

Al equipo de Penta Analytics, quienes desde un comienzo me brindaron su apoyo, entendimiento y tiempo cada vez que mi trabajo en la memoria lo requirió.

Finalmente, sólo me queda agradecer al pilar fundamental de mi vida: mi familia. A mi tía Elizabeth, que me recibió cuando me vine a estudiar al Nacional y me acogió en su casa como un hijo más, compartiendo conmigo mis aspiraciones y mis sueños.

A mi Madre, sin la cual no sería nada de lo que soy ahora, quien con su ejemplo de vida ha guiado mis pasos y a pesar de las distancias siempre ha estado en los momentos que la he requerido.

A mi hermana, que la he vuelto a descubrir en este último tiempo y que a pesar de mi manera de ser media arisca, siempre aprovecha cualquier oportunidad para demostrarme su cariño, sobre todo en los momentos difíciles que nos ha tocado sortear.

Sé que me queda mucha gente afuera, a todos ellos simplemente gracias por contribuir con un granito de arena a este trabajo. Este término de etapa, sólo simboliza el comienzo de una nueva, donde espero contar con mi gente para ayudarme a construir esta nueva etapa, la vida.

## Índice de Contenidos

1.- Antecedentes Generales.....	1
1.1 Historia.....	1
1.2 Transacciones con Tarjetas de Crédito.....	3
1.3 Regulación Vigente.....	8
2.- Objetivos y Alcance.....	9
2.1 Objetivo General.....	9
2.2 Objetivos Específicos.....	9
2.3 Alcances.....	9
3.- Marco teórico.....	10
3.1 Introducción al Manejo de Riesgo Crediticio.....	10
3.1.1 Riesgo Operacional.....	10
3.1.2 Riesgo de Mercado.....	10
3.1.3 Riesgo de Liquidez.....	10
3.1.4 Riesgo Crediticio.....	11
3.2 Tipos de Modelos de Medición.....	12
3.2.1 Modelos Cualitativos.....	12
3.2.2 Modelos Cuantitativos.....	12
3.3 Score de Comportamiento.....	14
4.- Metodología.....	18
4.1 Datos Disponibles.....	19
4.2 Selección de los Datos.....	20
4.3 Procesamiento de los Datos.....	20
4.4 Construcción de las Variables.....	21
4.5 Análisis Descriptivo.....	21
4.6 Construcción de Criterios Discriminantes y Selección de Variables.....	22
4.6.1 Carga de Datos y Selección de Variable Objetivo.....	22
4.6.2 Selección de la Muestra.....	23
4.6.3 Modelamiento de los Datos.....	23
4.7 Construcción y Testeo de Modelo de Prueba.....	23
4.8 Análisis.....	24
4.9 Conclusiones.....	24
4.10.- Definiciones Previas.....	24
4.10.1 Datos Disponibles.....	24
4.10.2 Determinación de Variables.....	25
4.10.3 Determinación Conjunto de Clientes.....	26
4.10.4 Construcción de las Variables.....	26
5.- Resultados.....	27
5.1 Base de Clientes.....	27
5.2 Variables Seleccionadas.....	29
5.3 Resultados Modelamiento.....	35
5.3.1 Resultados Modelo de Pruebas N°1.....	36

5.3.2 Resultados Modelo de Prueba N°2.....	37
5.3.4 Resultados Modelo Mixto N°1.....	40
5.3.5 Resultados Modelo Mixto N°2.....	42
5.4- Análisis de Resultados.....	44
6.- Conclusiones y Recomendaciones .....	48
7.- Bibliografía y Fuentes de Información.....	50
Anexos N°1 “Obligaciones Emisor-Titular Tarjetas de Crédito”.....	51
Anexos N°2 “WOE Variables de Muestra”.....	52
Anexos N°3 “Puntajes por variable” .....	55
Anexos N°4 “Listado Puntaje Final Obtenido” .....	59
Anexos N°5 “Matriz de Correlaciones” .....	69
Anexos N°6 “Medidas de Bondad Conjunta” .....	73

## Índice de Tablas

Tabla N° 1 "Cantidad de Tarjetas No Bancarias Vigentes al año 2011" .....	3
Tabla N° 2 "Cantidad de Tarjetas Bancarias Vigentes al año 2011" .....	4
Tabla N° 3 "Número Transacciones con tarjetas de crédito no Bancarias" .....	5
Tabla N° 4 "Porcentaje de Cartera al Día" .....	6
Tabla N° 5 "Porcentaje de Cartera de Clientes que se Encuentra en Mora" .....	7
Tabla N° 6 “Porcentaje de Cartera con Riesgo de cesación” .....	7
Tabla N° 7 "Número de Casos Modelo" .....	28
Tabla N° 8 "WOE CHAR_216" .....	30
Tabla N° 9 "WOE CHAR_234" .....	30
Tabla N° 10 "WOE CHAR_224" .....	31
Tabla N° 11 "WOE CHAR_222" .....	32
Tabla N° 12 "WOE CHAR_230" .....	33
Tabla N° 13 "WOE CHAR_232" .....	33
Tabla N° 14 "Listado Resumen Variables Seleccionadas" .....	34
Tabla N° 15 "Resultados Modelo de Pruebas N°1" .....	36
Tabla N° 16 "Resultados Clasificación Modelo de Pruebas N°1" .....	37
Tabla N° 17 "Resultado Modelo de Pruebas N°2" .....	38
Tabla N° 18 "Resultados Clasificación Modelo de Prueba N°2" .....	38
Tabla N° 19 "Listado Final Variables".....	40
Tabla N° 20 "Resultados Modelo Mixto N°1" .....	41
Tabla N° 21 "Resultados Clasificación Modelo Hibrido N°1" .....	41
Tabla N° 22 "Resultados Modelo Mixto N°2".....	43
Tabla N° 23 "Resultados Clasificación Modelo Mixto N°2" .....	43
Tabla N° 24 "Indicador K-S Modelo de Prueba" .....	46
Tabla N° 25 "Indicador K-S Modelo Mixto" .....	47
Tabla N° 26 "Indicador AUC".....	47

Tabla N° 27 "WOE CHAR_218" .....	52
Tabla N° 28 "WOE CHAR_228" .....	52
Tabla N° 29 "WOE CHAR_220" .....	53
Tabla N° 30 "WOE CHAR_226" .....	54
Tabla N° 31 "Puntajes Modelo de Prueba N°1" .....	55
Tabla N° 32 "Puntajes Modelo de Pruebas N°2" .....	56
Tabla N° 33 "Puntajes Modelo Mixto N°1" .....	57
Tabla N° 34 "Puntajes Modelo Mixto N°2" .....	58
Tabla N° 35 "Puntaje Final Modelo de Pruebas N°1" .....	59
Tabla N° 36 "Puntaje Final Modelo de Pruebas N°2" .....	61
Tabla N° 37 "puntaje Final Modelo Mixto N°1" .....	63
Tabla N° 38 "Puntaje Final Modelo Mixto N°2" .....	66
Tabla N° 39 "Matriz de Correlaciones Modelo de pruebas N°1" .....	69
Tabla N° 40 "Matriz de Correlaciones Modelo de Prueba N°2" .....	70
Tabla N° 41 "Matriz de correlaciones Modelo Mixto N°1" .....	71
Tabla N° 42 "Matriz de Correlaciones Modelo Mixto N°2" .....	72
Tabla N° 43 "Medidas De Bondad Modelo de Prueba N°1" .....	73
Tabla N° 44 "Medidas de Bondad Modelo de Pruebas N°2" .....	73
Tabla N° 45 "Medidas de Bondad Modelo Mixto N°1" .....	74
Tabla N° 46 "Medida de Bondad Modelo Mixto N°2" .....	74

### Índice de ecuaciones

Ecuación N° 1 "Pérdida Esperada" .....	11
Ecuación N° 2 "Valor de la Información" .....	15
Ecuación N° 3 " <i>Weight of Evidence</i> " .....	16
Ecuación N° 4 "Ecuación Modelo <i>Logit</i> " .....	16
Ecuación N° 5 " <i>Logit</i> Exponencial" .....	17
Ecuación N° 6 "Probabilidad de Ocurrencia de un Evento" .....	17
Ecuación N° 7 "Función de Probabilidades de la Muestra" .....	17
Ecuación N° 8 "Ecuación Log-Likelihood" .....	17
Ecuación N° 9 "Máxima Verosimilitud" .....	18
Ecuación N° 10 "Cálculo de Puntaje" .....	18

## Índice de Gráficos

Gráfico N° 1 "WOE CHAR_216" .....	30
Gráfico N° 2 "WOE CHAR_234" .....	31
Gráfico N° 3 "WOE CHAR_224" .....	31
Gráfico N° 4 "WOE CHAR_222" .....	32
Gráfico N° 5 "WOE CHAR_230" .....	33
Gráfico N° 6 "WOE CHAR_232" .....	34
Gráfico N° 7 "Distribución Acumulada Modelo de Pruebas N°1" .....	37
Gráfico N° 8 "Distribución Acumulada Modelo de Pruebas N°2" .....	39
Gráfico N° 9 "Distribución Acumulada Modelo Mixto N°1" .....	42
Gráfico N° 10 "Distribución Acumulada Modelo Mixto N°2" .....	44
Gráfico N° 11 "WOE CHAR_218" .....	52
Gráfico N° 12 "WOE CHAR_228" .....	53
Gráfico N° 13 "WOE CHAR_220" .....	53
Gráfico N° 14 "WOE CHAR_226" .....	54

## 1.- Antecedentes Generales

### 1.1 Historia

El tema a tratar en este trabajo de título corresponde al desarrollo de una estrategia de identificación y posterior incorporación de variables transaccionales dentro del modelo de clasificación de riesgo utilizado por la empresa Ripley S.A. Para estos efectos, se utilizará la información provista por la compañía relacionada con las transacciones realizadas a través de su tarjeta de crédito.

Ripley S.A es una empresa ligada a la industria del *Retail*. Se dedica a la venta al detalle de vestuario, accesorios y productos para el hogar a través del formato de tiendas por departamento. Por otro lado, la empresa está ligada al negocio financiero a través de la entrega de crédito directo a sus clientes mediante el uso de la tarjeta Ripley<sup>1</sup>.

La industria de *Retail* pertenece al sector “comercio, hotelería y restaurantes” de la economía. Este sector aporta cerca del 10% del PIB del país.<sup>2</sup> Las ventas totales de esta industria el año 2011 ascendieron a US\$ 40,3 MM<sup>3</sup>.

La historia de *Retail* en el mundo se remonta a principios del siglo XIX, época en que comienzan en Europa a aparecer las primeras tiendas por departamento. El objetivo de estas tiendas, más que satisfacer las necesidades, era “mostrar los productos en grandes salones de tiendas que parecían palacios”. Con el posterior desarrollo del modelo norteamericano de ventas, se derivó al diseño que las tiendas tienen en la actualidad.

En Chile, las primeras tiendas por departamentos se desarrollaron a principios del siglo XX siguiendo el modelo europeo de los grandes salones. Las primeras tiendas se ubicaron en el centro de la capital, específicamente en las calles Merced y Huérfanos, la primera tienda de estas características se fundó en 1910 y pertenecía a la cadena anglo-argentina Gath y Chávez. Con el correr de los años se fueron abriendo nuevas cadenas como Almacenes París, Falabella, Guendelman, Corona, entre otros.

La apertura de diferentes cadenas de negocios junto con el desarrollo económico que vivió el país durante el siglo pasado permitió la masificación de la industria a la vez

---

<sup>1</sup> Fuente: <http://www.ripley.cl/wcsstore/Ripley/inversionistas/>

<sup>2</sup> Fuente: Producto Interno Bruto por Clase de Actividad Económica. Serie Anual Nominal. Departamento de Cuentas

<sup>3</sup> Fuente: Asociación Gremial de Supermercados.



que democratizó el consumo en los diferentes niveles de la población. Este hecho provocó que las tiendas diversificaran su oferta de productos, permitiendo que nuevos sectores de la población tuvieran acceso a estos productos aquí comercializados.

El ingreso de diferentes sectores de la población y la necesidad de facilitar la entrada de otros, obligó a la industria a buscar nuevas maneras que permitieran la compra de especies, es así como se introduce el crédito en la industria, siendo en la década de los 70 cuando las principales empresas del país crearon sus primeros departamentos para administrar los sistemas crediticios, en los años 80 es introducido en la industria el plástico como medio válido para realizar transacciones. El crédito funciona como una línea de dinero ficticia que puede ser utilizada en la compra de bienes y/o servicios en los diferentes establecimientos habilitados para ello (siendo éstas de la misma casa comercial u otros establecimientos asociados). Dada la naturaleza de la herramienta, esta debe contener elementos de seguridad que permitan al emisor resguardarse de potenciales fraudes.

Durante toda su historia, el desarrollo del mercado de los créditos en el *Retail* ha contribuido, no solo a diversificar las fuentes de crédito en los mercados financieros, sino también a profundizarlo sustantivamente y transformarlo en uno de los puntales del negocio debido a la mayor cantidad de ingresos producto del cobro en intereses. Esta transformación ha permitido a familias de menores ingresos el acceso expedito a bienes durables, logrando así, tener una mejor calidad de vida<sup>4</sup>.

El sistema de créditos introducido en la industria significó también la necesidad de desarrollar diferentes sistemas de control para la administración de los créditos otorgados. Los sistemas de control utilizados en la actualidad tienen un carácter predictivo y buscan, a partir de información histórica, modelar el comportamiento de pagos que tendrán los diferentes clientes. Los modelos son utilizados para generar clasificaciones numéricas que resumen el comportamiento crediticio de los diferentes consumidores. Los modelos utilizan un conjunto de variables que contienen información histórica sobre el comportamiento de pago de los clientes, así como también otras que contienen información sobre los montos involucrados en las transacciones.

El negocio crediticio implica una evolución de los riesgos de las empresas participantes en el retail, ya que se introducen elementos nuevos como la necesidad de capital para el financiamiento de este tipo de operaciones, el riesgo de no pago de los créditos otorgados y los costos por la cobranza asociada que se debe realizar para obtener una recaudación adecuada.

---

<sup>4</sup> Fuente: Libro "Retail en Chile". Cámara de Comercio de Santiago.

## 1.2 Transacciones con Tarjetas de Crédito

En la actualidad, el mercado financiero en el retail se encuentra altamente desarrollado, de acuerdo a cifras del Centro de estudios del *Retail* (CERET<sup>5</sup>), existen en el país más de 20 millones de Tarjetas de Crédito no bancarias en el país. Según el Capítulo III.J de las Normas Financieras del Banco Central, se entiende por Tarjeta de Crédito "Cualquiera tarjeta u otro documento, que permite a su Titular disponer de un crédito otorgado por su Emisor y es utilizado por su Titular o Usuario en la adquisición de bienes o en el pago de servicios, vendidos o prestados por establecimientos afiliados al correspondiente sistema; sin perjuicio de las prestaciones adicionales que se podrán otorgar al Titular de la Tarjeta"<sup>6</sup>.

El siguiente cuadro muestra el número total de tarjetas de créditos no bancarias de las diferentes casas comerciales, el número corresponde al total de tarjetas que se encuentran vigentes<sup>7</sup> al cierre del periodo de información. También se indica el número de tarjetas que estando vigentes presentaron operaciones durante el año:

**Tabla N° 1 "Cantidad de Tarjetas No Bancarias Vigentes al año 2011"**

Emisor	Marca	Tarjetas vigentes al final del período informado	Tarjetas con operaciones en el período
ABC Inversiones Ltda.	Tarjeta ABC	141.597	23.143
Créditos Organización y Finanzas S.A. (Cofisa S.A.)	Tarjeta DIN y ABCDIN	1.070.421	290.561
Total ABC Inversiones Ltda. y Cofisa S.A.		1.212.018	313.704
Cencosud Administradora de Tarjetas S.A.	Jumbo	705.964	305.496
	Más Easy	223.520	70.813
	Más Paris	1.797.638	983.609
	Tur Bus Card	-	-
Total Cencosud Administradora de Tarjetas S.A.		2.727.122	1.359.918
Comercializadora y Administradora de Tarjetas Extra S.A. (1)	Xtra	-	-
Car S.A.	RIPLEY	2.687.674	1.098.009
	Mastercard CAR	53	33
Total Car S.A.		2.687.727	1.098.042
Consortio Tarjetas de Créditos S.A.	Visa Consortio	12.480	2.201
Administradora TMO S.A. (2)	Johnson's Multiopción	1.017.289	155.452
Inversiones SCG S.A. (3)	Tarjeta La Polar	N/D	N/D
Promotora CMR Falabella S.A.	CMR Falabella	3.900.044	910.869
	Visa Promotora CMR Falabella	1.717.316	1.122.762
Total Promotora CMR Falabella S.A.		5.617.360	2.033.631
Servicios y Administración de Créditos Comerciales Presto S.A.	PRESTO	1.403.041	468.556
<b>Total general</b>		<b>14.677.037</b>	<b>5.431.504</b>

Fuente: SBIF

Fuente: SBIF

<sup>5</sup> Fuente: Centro de Estudios del Retail. Departamento Ingeniería Industrial Universidad de Chile

<sup>6</sup> Artículo N°2 Disposiciones Generales Normas Financieras del Banco Central Capítulo III.J.1-1

<sup>7</sup> Tarjetas Vigentes: Tarjetas que a diciembre de 2011 se encuentran habilitadas para realizar operaciones. Fuente: SBIF

Según cifras oficiales, existen en el país alrededor de 4 millones de hogares<sup>8</sup>, esto implica que existen en promedio 3,7 tarjetas vigentes por cada hogar en Chile.

A nivel de tiendas por departamento Ripley se encuentra en el segundo lugar en cuanto a número de tarjetas vigentes y tarjetas con operaciones. Considerando los datos a nivel de *holding*, Ripley se ubica como el tercer actor de la industria en esta categoría.

Para completar el mapa general del uso de las tarjetas de crédito en nuestro país se adjunta la siguiente tabla que incluye la información a diciembre de 2011 con la cantidad y las transacciones que se realizaron utilizando tarjetas de crédito Bancarias.

**Tabla N° 2 "Cantidad de Tarjetas Bancarias Vigentes al año 2011"**

Emisor	Número de tarjetas vigentes al final del período		Total de tarjetas vigentes
	Titulares	Adicionales	
BANCO BICE	24.977	5.192	30.169
BANCO BILBAO VIZCAYA ARGENTARIA, CHILE (BBVA)	189.931	72.917	262.848
BANCO DE CHILE	1.145.853	178.012	1.323.865
BANCO DE CREDITO E INVERSIONES	550.115	183.065	733.180
BANCO DEL ESTADO DE CHILE	799.006	16.113	815.119
BANCO FALABELLA	69.844	2.332	72.176
BANCO ITAÚ CHILE	88.974	24.581	113.555
BANCO RIPLEY	20.652	3.465	24.117
BANCO SANTANDER-CHILE	1.545.471	321.112	1.866.583
BANCO SECURITY	37.009	6.373	43.382
BANCO PARIS	8.530	477	9.007
COOPEUCH	47.229	1.946	49.175
CORPBANCA	104.852	40.152	145.004
SCOTIABANK CHILE	129.010	21.058	150.068
BANCO INTERNACIONAL	2.177	121	2.298
HSBC BANK (CHILE)	1.404	211	1.615
SMU CORP S.A.	45.983	782	46.765
<b>Total general</b>	<b>4.811.017</b>	<b>877.909</b>	<b>5.688.926</b>

Fuente: SBIF

Fuente: SBIF.

De las tablas anteriores se aprecia que el número de tarjetas de crédito no bancarias casi triplica el número de las tarjetas Bancarias, a esto se debe sumar que según datos de la SBIF, el monto de las transacciones con tarjetas no bancarias casi

<sup>8</sup> Fuente: Instituto Nacional de Estadísticas.

duplica al de las transacciones con tarjetas bancarias, estos datos permiten entender el impacto que tiene en el país el uso de tarjetas de crédito de este tipo.

Para tener una idea del número total de transacciones que se realizan utilizando tarjetas de créditos como medio de pago, es que se adjunta la siguiente tabla, donde se indica el número total de transacciones que se realizaron con tarjetas no bancarias en el año 2011, así como también se indican los montos involucrados en las transacciones.

**Tabla N° 3 "Número Transacciones con tarjetas de crédito no Bancarias"**

Empresa	Acumulado anual	
	Número Transacciones	Monto (MM\$)
Tarjeta ABC	340.112	25.415
Tarjeta DIN y ABCDIN	4.709.618	226.932
<b>RIPLEY</b>	<b>29.159.952</b>	<b>595.096</b>
Mastercard CAR	83	13
Jumbo	10.258.290	290.802
Más Easy	1.651.204	49.295
Más Paris	22.972.672	703.804
Tur Bus Card	-	-
Xtra (2)	81	0,0701
Visa Consorcio	63.964	1.526
Johnson's Multiopción	2.892.044	58.567
Tarjeta La Polar (4)	N/D	N/D
CMR Falabella	36.427.279	980.319
Visa Promotora CMR Falabella	60.647.543	1.490.636
PRESTO	32.214.670	625.608
<b>Total general</b>	<b>201.337.512</b>	<b>5.048.014</b>

Fuente: SBIF

Fuente: SBIF.

De la tabla anterior, se desprende que en el año 2011, Ripley concentró cerca del 14,4% del total de transacciones realizadas y los montos involucrados correspondieron al 11.7% del total general.

Es conveniente también analizar la composición de las carteras de deudas generadas a partir de las transacciones realizadas con tarjetas de créditos, los datos corresponden al año 2011 y están expresados en términos porcentuales por trimestre.

**Tabla N° 4 "Porcentaje de Cartera al Día"**

Marca	Al día			
	Mar	Jun	Sep	Dic
CMR Falabella	87,4%	90,0%	90,2%	89,4%
Visa Promotora CMR Falabella	97,1%	96,6%	95,4%	94,4%
Johnson´s Multiopción	90,7%	91,8%	91,2%	90,8%
Jumbo	95,2%	95,3%	94,8%	95,5%
Más Easy	93,4%	93,3%	92,6%	93,0%
Más Paris	92,2%	92,3%	91,9%	92,9%
PRESTO	91,2%	92,1%	91,4%	91,9%
<b>RIPLEY</b>	<b>95,5%</b>	<b>94,9%</b>	<b>95,0%</b>	<b>95,1%</b>
Mastercard CAR	-	-	-	100%
Tarjeta ABC	92,6%	93,1%	93,0%	93,7%
Tarjeta DIN y ABCDIN	88,2%	87,7%	86,6%	86,9%
Tarjeta La Polar *	S/I	S/I	S/I	N/D
Visa Consorcio	96,2%	96,6%	97,1%	97,2%
<b>Total general</b>	<b>92,5%</b>	<b>93,2%</b>	<b>92,8%</b>	<b>92,7%</b>

Fuente: SBIF

**Fuente: SBIF.**

La siguiente tabla indica el porcentaje de créditos en mora<sup>9</sup> por trimestres que tuvieron las diferentes compañías<sup>10</sup>:

<sup>9</sup> Mora: Según la SBIF, por mora se entiende aquel crédito que tiene un pacto exigible cuyo plazo de vencimiento ya se ha cumplido.

<sup>10</sup> \*La información correspondiente a La Polar no se encuentra disponible debido a que se encuentra sujeta a revisión por los problemas que tuvo con el caso de las repactaciones unilaterales

**Tabla N° 5 "Porcentaje de Cartera de Clientes que se Encuentra en Mora"**

Marca	Total Moroso			
	Mar	Jun	Sep	Dic
CMR Falabella	12,6%	10,0%	9,8%	10,6%
Visa Promotora CMR Falabella	2,9%	3,4%	4,6%	5,6%
Johnson´s Multiopción	9,3%	8,2%	8,8%	9,2%
Jumbo	4,8%	4,7%	5,2%	4,5%
Más Easy	6,6%	6,7%	7,4%	7,0%
Más Paris	7,8%	7,7%	8,1%	7,1%
PRESTO	8,8%	7,9%	8,6%	8,1%
<b>RIPLEY</b>	<b>4,5%</b>	<b>5,1%</b>	<b>5,0%</b>	<b>4,9%</b>
Mastercard CAR	-	-	-	-
Tarjeta ABC	7,4%	6,9%	7,0%	6,3%
Tarjeta DIN y ABCDIN	11,8%	12,3%	13,4%	13,1%
Tarjeta La Polar *	S/I	S/I	S/I	ND
Visa Consorcio	3,8%	3,4%	2,9%	2,8%
Total general	7,5%	6,8%	7,2%	7,3%

Fuente: SBIF

Fuente: SBIF.

Del total de la cartera de morosos, la siguiente tabla muestra a nivel desagregado el porcentaje de clientes con potencial de entrar en cesación según las convenciones internacionales:

**Tabla N° 6 "Porcentaje de Cartera con Riesgo de cesación"**

Marca	30 días o más, pero menos de 90 días				90 días o más, pero menos de 180 días				180 días o más, pero menos de un año			
	Mar	Jun	Sep	Dic	Mar	Jun	Sep	Dic	Mar	Jun	Sep	Dic
CMR Falabella	2,3%	2,0%	2,2%	2,1%	0,8%	0,7%	0,7%	0,8%	5,0%	4,5%	4,0%	3,5%
Visa Promotora CMR Falabella	0,7%	0,9%	1,2%	1,2%	0,2%	0,3%	0,4%	0,4%	0,7%	1,2%	1,8%	2,3%
Johnson´s Multiopción	3,1%	3,1%	3,4%	3,3%	1,4%	1,7%	1,8%	1,7%	-	0,2%	0,2%	-
Jumbo	1,6%	1,5%	1,8%	1,6%	0,7%	0,7%	0,8%	0,8%	0,6%	0,6%	0,6%	0,6%
Más Easy	2,4%	2,5%	2,8%	2,7%	1,1%	1,2%	1,4%	1,4%	0,5%	0,4%	0,4%	0,4%
Más Paris	1,9%	1,9%	2,1%	2,0%	0,8%	0,9%	1,0%	1,0%	2,7%	2,4%	2,4%	2,0%
PRESTO	0,7%	0,7%	0,8%	0,7%	6,7%	5,9%	6,5%	6,3%	-	-	-	-
<b>RIPLEY</b>	<b>1,6%</b>	<b>1,8%</b>	<b>2,0%</b>	<b>1,6%</b>	<b>0,8%</b>	<b>1,0%</b>	<b>1,0%</b>	<b>1,0%</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>
Mastercard CAR	-	-	-	-	-	-	-	-	-	-	-	-
Tarjeta ABC	2,2%	1,5%	1,7%	1,6%	1,6%	1,8%	1,6%	1,6%	1,4%	1,4%	1,5%	1,5%
Tarjeta DIN y ABCDIN	3,4%	3,0%	3,4%	3,4%	2,8%	3,2%	3,5%	3,6%	2,5%	2,6%	3,0%	3,3%
Tarjeta La Polar *	S/I	S/I	S/I	ND	S/I	S/I	S/I	ND	S/I	S/I	S/I	ND
Visa Consorcio	0,1%	0,2%	0,2%	0,1%	1,1%	1,0%	0,7%	0,9%	2,2%	1,8%	1,7%	1,5%
Total general	<b>1,7%</b>	<b>1,6%</b>	<b>1,8%</b>	<b>1,6%</b>	<b>1,6%</b>	<b>1,6%</b>	<b>1,7%</b>	<b>1,7%</b>	<b>1,9%</b>	<b>1,6%</b>	<b>1,7%</b>	<b>1,7%</b>

Fuente: SBIF

Fuente: SBIF.

Estos datos puestos a modo de introducción, otorgan el contexto de la industria respecto al uso de las tarjetas de crédito y los datos de las carteras de deuda.

### 1.3 Regulación Vigente

La emisión y operación de tarjetas de crédito bancarias se encuentra sujeta a la regulación y vigilancia de la Superintendencia de Bancos e Instituciones Financieras y el Banco Central. Esta regulación se encuentra descrita en el Capítulo III.J del compendio de Normas Financieras del Banco Central.

La normativa obliga a las instituciones financieras que deseen operar tarjetas de crédito a inscribirse en el Registro de Emisores y Operadores de Tarjetas de Crédito, también indica que sólo las instituciones que tengan ventas anuales por sobre el millón de UF deben inscribirse en este registro.

Desde la perspectiva del emisor, la regulación vigente controla los niveles de capital y reserva, estableciendo que este no debe ser inferior a las 200.000 Unidades de Fomento. El límite de endeudamiento se establece en que no puede exceder 12,5 veces el capital pagado<sup>11</sup>. Por otro lado, la regulación obliga a las empresas que emitan tarjetas de créditos a someterse a un sistema de evaluación de gestión encargado a una empresa auditora externa. Finalmente, la regulación obliga a las empresas a informar sobre “cualquier hecho que revista el carácter de esencial conforme a los artículos 9°<sup>12</sup> y 10° de la ley N°18045 sobre Mercado de Capitales”<sup>13</sup>.

Este capítulo de la Norma Financiera también establece la definición de monto total de pago, el cual se define como “La suma consolidada de las adquisiciones de bienes y de los pagos de servicios y otras prestaciones, efectuados por los Titulares o Usuarios de las Tarjetas emitidas por la Empresa Emisora en los doce meses previos. Para efectos de esta normativa también se incluirán los avances en efectivo”.

Para el caso de los usuarios de las Tarjetas de Crédito, la regulación establece que debe existir una información mínima que permita una correcta identificación del emisor y el titular de la tarjeta. Se establecen también obligaciones para asegurar la transparencia del cobro de las tarjetas en cuanto a las comisiones, intereses y la mantención de la cuenta. En el anexo N°1 se adjunta una lista con los principales aspectos regulados entre emisor y titular de las tarjetas de crédito.

---

<sup>11</sup> Punto C, Apartado N°2 de III Empresas relacionadas, Capítulo III.J.1 “Emisión u Operación de Tarjetas de Crédito Banco Central.

\*La información correspondiente a La Polar no se encuentra disponible debido a que se encuentra sujeta a revisión por los problemas que tuvo con el caso de las repactaciones Unilaterales

<sup>12</sup> Artículo en modificación luego del caso La Polar.

<sup>13</sup> Sección f del apartado 2 del punto B de III Empresas relacionadas, Capítulo III.J.1 “Emisión u Operación de Tarjetas de Crédito Banco Central.

## **2.- Objetivos y Alcance**

### **2.1 Objetivo General**

Analizar un conjunto de datos transaccionales de la empresa para la generación de variables, que permitan identificar y cuantificar el impacto de la adición de este tipo de variables en un modelo de gestión de riesgo crediticio.

### **2.2 Objetivos Específicos**

- Construir una Base de Datos adecuada, con datos claros transparentes, limpios y ordenados, que permita la fácil elaboración de las variables necesarias para el desarrollo del modelo.
- Elaborar una metodología que permita establecer el correcto uso de los datos y los patrones de comportamiento.
- Obtener una mejora de a lo menos un 10% en los indicadores de asertividad del modelo, como el test Kolmogorov – Smirnov.<sup>14</sup>
- Realizar una comparación del poder predictivo de los modelos considerando la clasificación de clientes utilizada por la compañía y la clasificación elaborada especialmente para este trabajo.

### **2.3 Alcances**

La memoria descrita en este informe se enfoca principalmente en el análisis de los datos transaccionales de la empresa. Las transacciones analizadas corresponden a las efectuadas mediante el uso de la tarjeta de crédito de la empresa como medio de pago, tanto para compras realizadas en tienda como para aquellas realizadas en comercios asociados.

El trabajo sólo contempla la identificación de variables y la evaluación en un modelo preliminar, en ningún caso implica la confección de un modelo crediticio para la compañía.

El resultado final de esta memoria espera comprobar la hipótesis que las variables transaccionales aportan información relevante al modelo crediticio ya utilizado. Por otro lado se espera identificar las variables más relevantes para ser incorporadas al modelo crediticio ya utilizado y cuantificar su impacto.

---

<sup>14</sup> Indicador K-S: Indicador Kolmogorov-Smirnov que mide la bondad de ajuste entre dos distribuciones de probabilidad.



Se espera que con la identificación de las variables, el modelo actual usado por la compañía mejore su indicador K-S en al menos un 10%.

### **3.- Marco teórico**

Para la realización de este trabajo, es necesario presentar las referencias conceptuales sobre las que se sustenta la investigación. Para ello se comenzará por definir qué se entiende por riesgo y cuáles son los principales tipos de riesgo aplicados al ámbito financiero que existen. Posteriormente se profundizará en la descripción del tipo de riesgo específico sobre el que se desarrollará el trabajo para finalmente abordar los elementos teóricos que permiten su medición y las distintas herramientas que se han desarrollado para este propósito.

#### **3.1 Introducción al Manejo de Riesgo Crediticio**

En términos generales, el riesgo se entiende como la posibilidad de que un suceso inesperado altere los resultados que se esperan de un proceso determinado. La sola probabilidad de que un suceso tenga un resultado distinto al esperado genera incertidumbre en los individuos relacionados y obliga a las distintas partes involucradas a buscar mecanismos para la medición y gestión de éstos.

Desde el punto de vista financiero, existen cuatro clasificaciones principales de riesgo:

##### **3.1.1 Riesgo Operacional**

Se entiende la pérdida que surge cuando procesos internos de las empresas se han realizado de manera inadecuada. Pudiéndose deber a fraudes realizados por operadores, fallas en los sistemas involucrados en los negocios o problemas derivados del daño a activos físicos de las empresas.

##### **3.1.2 Riesgo de Mercado**

Se refiere al riesgo en el que están inmersas todas las entidades que interactúan en un determinado mercado. Se debe a cambios en las tasas de intereses decretadas por el Banco Central o movimientos en el precio de las acciones.

##### **3.1.3 Riesgo de Liquidez**

Ocurre cuando existen dificultades para solventar aumentos en los activos o cuando una institución no es capaz de cumplir con sus obligaciones.

### 3.1.4 Riesgo Crediticio

Surge de la posibilidad que ocurra un incumplimiento de la contraparte cuando el valor de un contrato con una institución financiera es positivo.<sup>15</sup> En otras palabras, es la posibilidad de que la contraparte de una institución no cumpla las obligaciones que tiene pactadas con ella.

Debido a la existencia de este tipo de riesgo, se hace necesario el desarrollo de técnicas que permitan a las diferentes entidades determinar los niveles de solvencia de los clientes sujetos a préstamos y con ello gestionar de mejor manera sus diferentes carteras de clientes.

Cuando se quiere medir el riesgo de crédito se deben tomar en cuenta los siguientes factores:

- Probabilidad de Incumplimiento: Es una medida de que tan probable es que un cliente incumpla con las obligaciones pactadas.
- Exposición de la cartera: Es una medida de la cantidad adeudada al momento de producirse el incumplimiento.
- Tasa de Recuperación: Se define como la porción de la exposición que es recuperada en caso de producirse un incumplimiento.

Con estos tres elementos se puede calcular la pérdida esperada de una cartera de créditos, la cual queda definida de la siguiente manera:

$$Pérdida Esperada = EAD * PI * (1 - r)$$

#### Ecuación N° 1 "Pérdida Esperada"

Esta pérdida puede ser entendida como lo que en promedio puede perder una cartera en un horizonte determinado, dada una cierta distribución de las pérdidas. Permite estimar la reserva preventiva que tiene que tener una entidad contra riesgos crediticios.

Por otro lado, se define la pérdida inesperada como aquella que ocurre cuando las pérdidas efectivas superan a las estimadas, este valor corresponde a la diferencia entre ambos valores. La pérdida inesperada se debe principalmente a volatilidades en la exposición, a las probabilidades de incumplimiento y también a las correlaciones

---

<sup>15</sup> Fuente: HULL, John C. Options, Interest Rates Markets. En Futures and Others Derivatives 5ta Edición EE.UU. Prentice Hall, 2003 pp. 93-124.

existentes entre los incumplimientos, que provocan que las pérdidas sean mayores a las que se esperan.

### **3.2 Tipos de Modelos de Medición**

La existencia de este tipo de riesgos obliga a la búsqueda, desarrollo y aplicación de distintos sistemas y políticas para analizar y evaluar riesgos de manera de controlar las potenciales pérdidas, maximizando así las rentabilidades de las diferentes carteras.

La evolución que han tenido las distintas herramientas para mitigar los efectos producidos por este tipo de riesgos, han ido de la mano del desarrollo de mejores paquetes estadísticos que realizan complejos cálculos matemáticos en un menor tiempo y con una mayor precisión. El desafío de estos modelos ha estado centrado en minimizar las pérdidas que surgen cuando se está relacionado con el otorgamiento de créditos.

En un comienzo, el control del riesgo crediticio se realizaba mediante la implementación de una serie de cláusulas que restringían el otorgamiento de créditos así como también la exigencia de avales. Este tipo de restricciones se enmarcan dentro de los modelos cualitativos para la estimación del riesgo crediticio.

#### **3.2.1 Modelos Cualitativos**

Los modelos cualitativos tienen la característica que se utilizan una serie de criterios de carácter subjetivo, principalmente basados en el juicio de experto realizado por la persona encargada de evaluar. El modelo más conocido de estas características es el modelo de las cinco “C” que considera 5 factores:

- Capital involucrado: Mide el capital disponible de los individuos o empresas que solicitan un crédito.
- Capacidad de Pago: Mide la capacidad de pago del individuo, normalmente considera la volatilidad asociada a los ingresos del deudor.
- Colateral del Préstamo: Considera la “garantía” asociada al préstamo y las posibilidades de disminuir la pérdida en caso de que se incurra en incumplimiento.
- Carácter: Considera información relativa al prestatario, sirve para tener alguna referencia de la persona a la que se le está prestando.
- Condiciones: una variable exógena que describe el ambiente externo que involucra la emisión del bono.

#### **3.2.2 Modelos Cuantitativos**

Son aquellos que utilizan técnicas de modelamiento de la información para transformar los datos disponibles y con ello derivar las estimaciones de escenarios y

decisiones donde los individuos toman sus decisiones. Este tipo de modelos se divide entre modelos no paramétricos y modelos paramétricos.

Un modelo no paramétrico se caracteriza por no requerir de supuestos especiales sobre la distribución de los datos utilizados. Un ejemplo de modelo no paramétrico es el de redes neuronales, que mediante el uso de técnicas econométricas, crea un modelo que emula una red de neuronas, capaz de asumir decisiones en torno a un riesgo de referencia.

Por otro lado, los modelos paramétricos son aquellos que asumen ciertas características dadas en los datos disponibles, las principales técnicas paramétricas utilizadas son el análisis discriminante, los modelos de probabilidad lineal y la regresión logística.

El trabajo a desarrollar considera la utilización de la metodología para la generación de *score* de clasificación. Esto se define como el uso de modelos estadísticos para transformar información relevante en indicadores numéricos que permitan guiar una decisión relacionada con el otorgamiento de créditos. La metodología del *score* de clasificación provee un gran valor cuando es usada para guiar una decisión que afecta a los clientes, ya que el emisor de un préstamo crea una serie de escenarios y en base a éstos, se definen diferentes políticas o criterios para la aprobación de un crédito, la extensión del monto máximo de cupo e incluso la tasa de interés a la que se debe prestar a un cliente.

Dependiendo de cómo es usada, esta metodología se clasifica en diferentes tipos<sup>16</sup>:

- *Application Score*: Es usado principalmente cuando se origina una petición de crédito. Utiliza información socio demográfica, comportamiento anterior y la información crediticia actual del cliente.
- *Behavioural Score*: Utilizado para el manejo de las cuentas, como el límite del crédito. Se centra en el manejo de una cuenta individual.
- *Collections Score*: Utilizado para identificar un número de buenos pagadores dentro de un grupo de riesgo con mayoría de malos pagadores.

---

<sup>16</sup> ANDERSON, Raymond. *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk, Management and Decision Automation*. Oxford: Oxford University Press, 2007. ISBN 978-0-19-922640-5 pp. 4 -26.

- *Bureau Score*: Resume el comportamiento crediticio de un cliente en particular presente en la cartera de crédito con un índice que indica la probabilidad de incumplimiento.

Un *score* funciona asignando puntajes a cada uno de los deudores de una determinada cartera, este puntaje es ordenado en un *ranking* que permite discriminar a los que son buenos y malos pagadores. Este tipo de modelos es de gran utilidad para la identificación de los clientes de acuerdo a su comportamiento, clarificando aquellos que están en condición de ver aumentado su cupo, identificando de manera temprana a los que presentan mal comportamiento y a aquellos a los que definitivamente se les debe cerrar la cuenta.

Esta metodología utiliza como variables una serie de características propias de los clientes, la inclusión de estas características se hace en una fase previa al estudio de las variables. Las características de los individuos que serán consideradas dependerán del tipo de modelo que se desee usar. La idea, independiente del tipo de *score* que se utilice, es clasificar al cliente en una determinada categoría a partir de la gran cantidad de información que se tiene disponible de él.

### **3.3 Score de Comportamiento**

La metodología específica que se utilizará en el desarrollo de la memoria está relacionada con *Behavioural Scoring*, modelo que utiliza variables que derivan del comportamiento de los consumidores para generar un *score* de clasificación, el tipo de variables que típicamente se incorporan dependen de:

- **Historial de Pago:** Variables que resumen el comportamiento de pago que los clientes han tenido en el periodo de tiempo analizado, pondera positivamente a los individuos que han tenido buen comportamiento de pago.
- **Deuda:** Se incorporan variables que incluyan información respecto al nivel de deuda que mantengan los individuos con las diferentes entidades.
- **Historial de Uso:** Corresponde a un conjunto de variables que describan la frecuencia de uso de tarjeta.

Esta metodología consiste en escoger un determinado tiempo  $t$ , desde donde se observa el comportamiento transaccional de los individuos en un periodo de tiempo pasado para construir una serie de variables independientes. La variable dependiente o explicada del modelo es la ocurrencia de un incumplimiento, esta variable es de carácter binario donde 1 corresponde ocurrencia del incumplimiento y 0 el caso contrario, una vez hecho este procedimiento se procede a realizar calibraciones para un pronóstico del comportamiento futuro y evaluar estas características.

Antes de definir cuáles serán las variables que se utilizarán, se debe definir cuáles serán los criterios para determinar cuándo un individuo es “Bueno” o cuando es “Malo”.

Un Bueno se define como aquella observación que tiene un resultado deseado<sup>17</sup>, mientras que un Malo será el resultado opuesto. Por resultado deseado se entenderá aquella situación crediticia que es aceptable para la compañía. La literatura tradicional considera a un Bueno como aquel individuo que nunca ha presentado incumplimiento en sus pagos o ha presentado un incumplimiento en un periodo menor a 30 días. Por otro lado, un Malo es aquel individuo que presenta una mora superior a 90 días, que haya incurrido en mora 2 veces en al menos 60 días o 3 o más veces en los últimos 30 días.

El modelo busca discriminar entre estas dos opciones y diferenciar de la mejor manera posible las características de estos tipos de individuos.

Para la determinación de las variables que se utilizarán para enriquecer el modelo, se trabajará con la base de datos provista por la empresa. Los criterios de selección considerarán elementos basados en la experiencia, donde el juicio experto se utilizará como uno de los principales criterios.

Una vez que se han definido las variables y se han extraído y analizado los datos, se procede a escoger alguna característica particular de los individuos que permita categorizar las variables en diferentes intervalos, y con ello medir la cantidad de Bueno/Malo presentes en cada tramo. Esta división se realiza determinando rangos de valores para cada variable analizada, dentro de los cuales se cuenta la cantidad de individuos Buenos y Malos presentes en ese rango. Esto permite definir el valor de la Información como:

$$VI = \sum_{i=1}^N (\# \text{Bueno}_i - \# \text{Malo}_i) * \ln \left( \frac{\# \text{Bueno}_i}{\# \text{Malo}_i} \right)$$

Ecuación N° 2 "Valor de la Información"

Donde:

# **Bueno<sub>i</sub>** : Corresponde a la proporción de Buenos del intervalo "i" sobre el total de Buenos.

# **Malo<sub>i</sub>** : Corresponde a la proporción de Malos del intervalo "i" sobre el total de Malos de la muestra

---

<sup>17</sup> ANDERSON, Raymond. *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk, Management and Decision Automation*. Oxford: Oxford University Press, 2007. ISBN 978-0-19-922640-5 pp. 4 -26.

El índice “N” corresponde a la cantidad de intervalos presente en cada variable analizada. Este indicador permite medir cuanta información aporta cada variable al modelo.

Basados en este indicador se considera lo siguiente<sup>18</sup>:

- Valor menor a 0,02 no tiene poder predictivo.
- Valor entre 0,02 y 0,1 es débil.
- Valor entre 0,1 y 0,3 intermedio.
- Valor mayor a 0,3 Fuerte poder predictor.

También se debe considerar el indicador conocido como “peso de la evidencia” o WOE.

$$\ln\left(\frac{\#Bueno_i}{\#Malo_i}\right) = WoE$$

**Ecuación N° 3 "Weight of Evidence"**

Este indicador mide la “fuerza” de un atributo e indica cuán representado en el atributo están los Buenos o Malos.

Una vez que se han determinado cuales son las variables pertinentes de acuerdo a los criterios recién definidos, se procede a elaborar el modelo de regresión con el cual se testeará finalmente el poder predictivo de las variables encontradas. La técnica que tradicionalmente utilizan los modelos de *scorecard* es la regresión logística, cuya ecuación se describe a continuación:

$$Logit(p) = \beta_0 + \sum \beta_i * x_i$$

**Ecuación N° 4 "Ecuación Modelo Logit"**

Donde:

P= Probabilidad de ocurrencia de incumplimiento

$\beta_0$ = Intercepto

$\beta_i$ = Parámetros estimados

---

<sup>18</sup> SIDDIQI, Naeem. Scorecard Development Process En *Designing Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring*. New Jersey: John Wiley & Sons, Inc., 2006. ISBN-13: 978-0-471-75451-0

$x_i$ = Variables

Otra forma de caracterizar este método es la siguiente:

$$\text{prop}(Y_i = j) = \frac{e^{\beta'_{kj}x_{ki}}}{\sum e^{\beta'_{kj}x_{ki}}}$$

**Ecuación N° 5 "Logit Exponencial"**

Esta técnica utiliza un conjunto de características para estimar una probabilidad de ocurrencia de un cierto evento, que en este caso se refiere a si el cliente está o no en categoría de incumplimiento.

*Logit* asume que la componente aleatoria es iid (independiente, idénticamente distribuida), por lo que el modelo captura variaciones sistemáticas de las preferencias, no aleatorias. También permite capturar patrones en elecciones repetidas con los factores observables que se incluyen en el modelo.

El Modelo permite estimar una respuesta binaria en base a la utilización de una función de transformación  $F(x)$  estrictamente creciente, cuyos intervalos son  $[0,1]$ , permitiendo la utilización del modelo para la estimación de probabilidades de ocurrencia

La estimación de los parámetros  $\beta$ 's del modelo se realiza mediante el método máxima verosimilitud, donde se tiene que las probabilidades de que el suceso ocurra o no son:

$$P(Y/x) = p^y * (1 - p)^{1-y}$$

**Ecuación N° 6 "Probabilidad de Ocurrencia de un Evento"**

Para la muestra se tiene que:

$$P(Y_1, Y_2, \dots, Y_N) = \prod_i p_i^{Y_i} * (1 - p_i)^{1-Y_i}$$

**Ecuación N° 7 "Función de Probabilidades de la Muestra"**

Tomando logaritmos

$$\text{Log}(P(Y)) = \sum_i y_i * \log\left(\frac{p_i}{1 - p_i}\right) + \sum_i \text{Log}(1 - p_i)$$

**Ecuación N° 8 "Ecuación Log-Likelihood"**

Se expresa la ecuación en función de los parámetros que se desea estimar:



$$L(B) = \sum_i y_i * x_i * b - \sum_i \text{Log}(1 + e^{x_i * b})$$

**Ecuación N° 9 "Máxima Verosimilitud"**

Desde donde se deriva para encontrar el máximo en función de un parámetro inicial que se debe estimar. Una diferencia sustancial con los modelos de regresión lineal, radica en el hecho de que los ponderadores resultantes no pueden ser interpretados de manera directa. Los ponderadores resultantes de la estimación por *Logit* deben ser interpretados mediante su variación marginal, donde un cambio en el valor de  $x_{ij}$  indica que ha ocurrido un cambio en la probabilidad de incumplimiento de un individuo.

A partir de la estimación realizada mediante el método de máxima verosimilitud, se determina la calidad de los individuos mediante la construcción de un puntaje ponderado que utiliza el valor de cada variable y los ponderadores estimados.

El puntaje que se asigna a cada variable, depende del coeficiente  $\beta$  y del *WOE* respectivo, es calculado para los diferentes tramos en los cuales se encuentra dividida la variable. A continuación se detalla la fórmula que permite obtener dicho cálculo:

$$Puntaje = \sum_{j,i=1}^{k,n} \left( - \left( woe_j * \beta_i + \frac{a}{n} \right) * factor + \frac{offset}{n} \right)$$

**Ecuación N° 10 "Cálculo de Puntaje"**

Donde:

*WOE*: "Peso de la evidencia" de cada variable.

$\beta$ : Coeficiente de la regresión de cada variable.

*n*: Número de variables incluidas en la regresión.

*k*: Número de atributos en cada variable.

*Factor*: Ponderador ligado al cociente de probabilidades de ocurrencia de un evento.

*Offset*: Factor relacionado con la escala de *score*.

*n*: Número de características.

Una vez que se ha definido el ordenamiento de acuerdo a los puntajes que cada individuo haya obtenido, se procede a determinar cuál es el punto de corte ad-hoc, para determinar si se otorgan o no un determinado crédito.

#### **4.- Metodología**

En la actualidad, los modelos crediticios usados en la industria utilizan variables que intentan describir el comportamiento de los clientes desde una perspectiva agregada de los datos. Estos modelos utilizan variables como evolución del monto de deuda o la cantidad de pagos realizados en un determinado periodo de tiempo, el

número de compras realizadas y algunas variables relacionadas con los niveles y características de facturación de los distintos clientes. Lo que pretende este estudio es analizar si la incorporación de variables que brinden una información más detallada del comportamiento de los clientes, al capturar patrones de conducta provenientes del tipo de productos adquiridos y los cambios registrados en estos patrones, permite enriquecer el modelo actualmente utilizado

La metodología para la construcción de las variables empleadas en este trabajo se basa en el proceso de extracción de datos conocido como KDD, este proceso permite la extracción de información sobre grandes volúmenes de datos y se define como “el proceso no trivial de identificación en los datos de patrones válidos, nuevos, potencialmente útiles y finalmente comprensibles”.<sup>19</sup>

En lo siguiente se describirá el proceso completo que fue desarrollado para efectos de esta memoria, donde se describen desde las fuentes iniciales disponibles para la obtención de información hasta los pasos seguidos para el desarrollo de los distintos modelos y regresiones estadísticas para la obtención de los resultados del trabajo.

#### **4.1 Datos Disponibles**

Los datos utilizados fueron proporcionados por la empresa Ripley S.A. Esta información corresponde a transacciones efectuadas por clientes en las diferentes tiendas de la empresa y aquellas que se realizaron en locales asociados, que involucran la utilización de la tarjeta de crédito de la compañía. Por local asociado se entenderá aquel local ajeno a Ripley donde puede ser utilizada la tarjeta de crédito de la tienda como medio de pago efectivo. El detalle de la información disponible para las transacciones realizadas en la tienda se encuentra a nivel de departamento, mientras que para comercios asociados esta se encuentra a nivel de rubro. La información puede ser desglosada de acuerdo a los montos involucrados en las transacciones, el número de cuotas pactadas, la cantidad de compras realizadas, fecha de las transacciones, tipos de productos comprados (si es que es un producto o un servicio) y otro tipo de información relevante para el manejo interno de la compañía como márgenes comerciales, financieros entre otros. La información se encuentra disponible en distintos repositorios internos de la compañía.

---

<sup>19</sup> Fuente: MOLINA, José Miguel y García, Jesús Técnicas de Análisis de Datos. Madrid. Universidad Carlos III, 2006.

## 4.2 Selección de los Datos

Esta fase del trabajo consistió en analizar y seleccionar las bases y los datos disponibles que finalmente se utilizaron. El análisis de datos se realiza para conocer la calidad y la cantidad de la información disponible. Este análisis se realizó considerando el horizonte de tiempo y los clientes incluidos como base para el desarrollo del modelo. El horizonte de tiempo es determinado mediante un juicio experto, el cual fue trabajado directamente con la empresa, esto se hace para aprovechar el conocimiento que tienen de su propio negocio.

El trabajo se desarrolló con toda la base de clientes con la que cuenta la compañía (aproximadamente cuatro millones de clientes), dentro de esta base se deben separar a los clientes que por distintos motivos son rechazados, mientras que el resto es clasificado como “Bueno”, “Malo” e Indeterminado.<sup>20</sup>

La selección final de los repositorios a utilizar fue determinada en conjunto con los miembros de la empresa y para ello se analizó el tipo de información que se deseaba testear. Finalmente se utilizaron bases con información de las transacciones en la tienda, de transacciones en comercios asociados, información adicional sobre productos especiales ofrecidos por la compañía.

## 4.3 Procesamiento de los Datos

Debido a la gran cantidad de información que deriva del proceso de venta de los diferentes bienes y servicios comercializados por la tienda, se hace necesaria la revisión de los datos disponibles. Este análisis tiene por objetivo la constatación de ciertas características que deben tener los datos para asegurar la confiabilidad de los resultados que se obtengan en el trabajo a realizar.

Las características que principalmente se consideran son:

**Número de casos y cantidad de variables por casos:** Se debe asegurar que existan un número suficiente de casos que permita la realización de un modelo predictivo, dichos casos deben contener una buena cantidad de variables para explicar el comportamiento que se desea modelar. Este trabajo contempla la utilización de la base de clientes de la empresa, la que contiene cerca de 700.000 casos para cada mes considerado (mayor detalle en la sección 4.10.3). Por casos se entenderá la información provista por cada cliente que se incluya en la muestra.

---

<sup>20</sup> Una muestra indeterminada es aquella que debido a su comportamiento no es posible clasificarla como Bueno o Malo, el detalle de la clasificación se encuentra en la sección 4.6.1

**Estacionalidad:** Debido a que los datos utilizados se generan a partir de la interacción comercial entre clientes y la compañía, y que estas interacciones están sujetas a variaciones propias de la naturaleza del mercado en el que están inmersas, es que muchos de los datos presentan variaciones cíclicas que deben ser consideradas en la construcción de las variables a utilizar y en el análisis de resultados.

**Consistencia:** La construcción de variables, la definición de Bueno/Malo y los periodos de tiempo que se utilizan para definir cuando un cliente entra en mora deben ser consistentes con los criterios utilizados en el actual modelo empleado por la compañía.

**Antigüedad de los Casos:** Los casos escogidos deben tener la antigüedad suficiente en la tienda para asegurar un número de observaciones que permita un buen nivel predictivo de las variables escogidas, esto se logra incorporando en el criterio de las exclusiones a aquellos clientes que hayan abierto una cuenta en un periodo de tiempo tal que no aporten información suficiente.

**Transformaciones Adiciones:** Con la primera obtención de los datos es necesario realizar un análisis sobre la calidad de estos y realizar las transformaciones necesarias en caso de tener datos nulos o sin información.

#### **4.4 Construcción de las Variables**

Esta fase consistió en la construcción de las 157 variables que se utilizaron en el modelo. Para la determinación y construcción de las variables se utilizaron herramientas provenientes de *Data Mining* y *Marketing*, las que permitieron por una parte la construcción de variables orientadas a capturar patrones de conducta que ayudan a predecir comportamiento y por otra la incorporación de elementos referidos al manejo de categorías de productos, los cuales permitirán eliminar efectos en las variables que puedan salir de campañas promocionales que emprenda la empresa.

#### **4.5 Análisis Descriptivo**

Una vez identificadas y construidas las variables a utilizar se procede a hacer un análisis descriptivo de éstas, con el fin de conocer su naturaleza, consistencia y la relación que tienen con las demás variables en cuanto a su correlación. Este análisis se realiza con el objetivo de corregir problemas o cambiar y agrupar variables que puedan restar poder predictivo al modelo. También se debe analizar el comportamiento de las variables con respecto al objetivo, que es saber si un cliente está dentro de la categoría de incumplimiento o no.

## 4.6 Construcción de Criterios Discriminantes y Selección de Variables

En esta parte del trabajo se utilizó una herramienta computacional facilitada por la empresa llamada “*Model Builder*”. Este *software*, desarrollado por FICO<sup>21</sup>, permite la construcción de un modelo predictivo a partir del procesamiento de la base de datos con las variables construidas para este estudio.

Si bien es cierto que este programa puede realizar de forma completa un *scorecard* de comportamiento, para efectos de este trabajo sólo se utilizó para determinar el rango de apertura que deben tener los intervalos en los que se agrupan los valores de las distintas variables y que permiten el posterior cálculo del WOE y del valor de la información (VI) de las variables seleccionadas.

El modelamiento de los datos se realizó en las siguientes etapas:

### 4.6.1 Carga de Datos y Selección de Variable Objetivo

En esta fase se carga la base de datos a utilizar y se crean las variables de desempeño con las que se trabajará en el modelamiento posterior. Las variables creadas son “*Target*” y “*Performance*”. La primera es la variable objetivo que marca a los Buenos y Malos, la segunda permite la identificación de los casos que serán excluidos del modelamiento. La variable objetivo identifica al cliente clasificado como malo. En esta memoria se trabajó con dos clasificaciones de clientes distintas, una que corresponde a la utilizada por la compañía y otra que considera los siguientes criterios:

Por cliente Malo se entenderá a aquel que presente en los 12 meses posteriores a cada uno de los meses que sirven de base una mora superior a 90 días en una oportunidad o, en el mismo periodo presente una mora superior a 60 días e inferior a los 90 en dos oportunidades o, que en al menos tres oportunidades haya presentado una mora superior a los 30 días e inferior a los 60 días”.

Por cliente Bueno, se entenderá a aquel que tenga siempre una cantidad inferior a los 30 días de mora en el periodo de tiempo analizado, mientras que un cliente estará clasificado como Indeterminado si presenta algún comportamiento intermedio a los descritos para clasificar a los clientes como Bueno o Malo.

---

<sup>21</sup> FICO: Empresa especializada en el desarrollo de herramientas para mejorar la gestión de empresas.

#### **4.6.2 Selección de la Muestra**

Una vez creada la variable objetivo, la siguiente etapa es la creación de la muestra estratificada sobre la cual se hará el modelamiento. Este proceso se hace mediante un selector que tiene el programa que permite seleccionar el número de casos a modelar y asignar el peso de la representación que tendrá la muestra. Para efectos de esta parte del trabajo se seleccionó una muestra de 200.000 casos (divididos en partes iguales entre *buenos y malos*), en la cual el 80% será utilizado para entrenar el modelo y el 20% restante para la validación. La selección de la cantidad de casos se acordó con la gente de la empresa buscando mantener los mismos criterios considerados por ellos en la elaboración de las aperturas para cada variable.

#### **4.6.3 Modelamiento de los Datos**

Con el número de casos y la variable objetivo definida, es posible realizar el modelamiento con el programa, el cual se realiza en 2 etapas. En la primera fase, el programa analiza las distintas variables a modelar y mediante algoritmos de clasificación, en donde considera la dispersión de los datos, propone los intervalos en los que deben ser agrupadas las distintas variables, entregando el número de buenos y malos correspondientes a cada intervalo. Una segunda fase permite realizar cambios sobre el resultado anterior del modelamiento, donde se modifican de manera “manual” los límites de los intervalos determinados por el programa, lo que permite corregir el WOE de aquellas variables que presenten inconsistencias en los valores inicialmente determinados. Una vez que se concluye la modificación, el programa reprocesa las variables recalculando el número de buenos y malos para cada intervalo. Las salidas que entrega el programa permiten calcular el valor de la información para cada una de las variables construidas y con ello seleccionar las que presenten un mayor poder predictivo.

#### **4.7 Construcción y Testeo de Modelo de Prueba**

Con las salidas obtenidas en el modelamiento utilizando el software facilitado por la empresa se construyó el modelo para testear las variables seleccionadas. Esta labor se realizó utilizando el programa estadístico SPSS, la técnica empleada fue la regresión logística sobre la variable binaria objetivo a modelar, la que corresponde a si es que el individuo es malo o bueno (1 si es malo, 0 si no lo es). La regresión se hizo sobre los WOE obtenidos a partir de las aperturas determinadas en la fase anterior. Para la selección de las variables finales del modelo se utilizó como criterio el valor de la Información, el cuál fue complementado con los resultados obtenidos a partir de una regresión logística utilizando el método de selección hacia adelante (*forward*) utilizando como criterio de selección la técnica de *Wald*, que incorpora variables paso a paso empezando por aquellas que tengan mayor significancia en cada paso hasta finalizar en

el punto en que el incremento en el coeficiente de determinación al incluir una nueva variable ya no es relevante.<sup>22</sup>

## **4.8 Análisis**

El análisis de los resultados se realiza para cada modelo desarrollado. Estos modelos difieren entre sí en la clasificación de clientes utilizada, lo que se traduce en diferencias en la cantidad de Buenos y Malos para cada modelo. Los resultados son analizados de manera independiente para cada modelo y se comparan con los resultados que se tienen del modelo utilizado por la empresa. Por otro lado, se analizan las diferencias entre las clasificaciones utilizadas y el impacto que estas diferencias tienen en el modelo desarrollado.

## **4.9 Conclusiones**

Finalmente se elaboran las conclusiones de la memoria en base a los resultados obtenidos y los objetivos inicialmente planificados. La principal conclusión que se espera obtener es confirmar que variables de carácter transaccional tienen impacto positivo dentro de un modelo de clasificación crediticia y permiten mejorar el poder predictivo del modelo.

## **4.10.- Definiciones Previas**

En la presente sección se describe el proceso de determinación de las variables que se utilizaron en el estudio, definiendo el periodo de tiempo sobre el cual se trabajó, así como también los criterios que fueron utilizados para la determinación de la base de clientes con la que se desarrolló la investigación.

### **4.10.1 Datos Disponibles**

El primer elemento a definir es el periodo de tiempo que se considerará. Para efectos de esta investigación se trabajó con datos del periodo comprendido entre el 01 de enero de 2009 y el 31 de diciembre de 2010. El periodo de tiempo considerado no logra abarcar un ciclo completo como lo recomienda el acuerdo de Basilea II, pero dado los criterios utilizados por la empresa se consideran suficientes para la realización de este trabajo.

Los datos disponibles para el trabajo contemplan la base de transacciones realizadas en la tienda, la base de transacciones realizadas en comercios asociados, la

---

<sup>22</sup> Fuente: Manual SPSS para Windows. Universidad Complutense de Madrid [En línea] <<http://www.ucm.es/info/dosis/Preventiva/doctorado/TEMA13.pdf>> [Consulta: 06 Agosto 2012]

base maestra de productos y la base maestra de comercios asociados. De las distintas bases se consolidó la información y se construyó el set final de variables utilizadas en este trabajo.

En relación a la información de clientes, se trabajó a partir de su identificador único, donde sólo se tuvo acceso a su comportamiento histórico de pagos y a la cantidad de días de atraso que pudo haber presentado un individuo, en los distintos meses observados.

#### **4.10.2 Determinación de Variables**

Como ya se ha mencionado, para este trabajo se construyeron 157 variables, éstas intentan describir distintos comportamientos que pueden ser capturados por medio de la información transaccional disponible.

Los productos involucrados en las transacciones se encuentran detallados a nivel de *SKU*<sup>23</sup>. Cada producto tiene asociado un código correspondiente a la división general a la que pertenece, desagregándose hasta nivel de marca y modelo de producto. En la medida en que se avanza en las divisiones de los productos, el árbol se va abriendo e incluye un mayor número de elementos. Para efectos de esta memoria se trabajó a nivel de área, la que se encuentra dividida en 30 sectores distintos. Esta decisión se trabajó con el equipo de la empresa y se tomó en base al juicio experto de los agentes sobre el negocio de la empresa.

El comportamiento que intentan capturar las variables se describe de la siguiente manera:

- Compras con la tarjeta en las diferentes sucursales de la tienda: la Frecuencia de compra, tiempo entre compras, montos de compra, cuotas pactadas, pago de cuotas entre otras.
- Compras realizadas en comercios asociados en los siguientes rubros: Supermercados, Farmacias, Bencineras, Ópticas y Locales de comida.
- Servicios financieros adquiridos: Incluyendo montos involucrados y número de cuotas pactadas.
- Departamentos en los que compran los clientes: Describe principalmente “la lealtad” a través de la frecuencia en los diferentes departamentos y el “*ticket promedio*”<sup>24</sup> asociado a las compras en estas áreas.

---

<sup>23</sup> SKU= Stock Keeping Unit. Es un identificador que permite dar seguimiento a u producto ofrecido por cada empresa. Este indicador es único para cada producto comercializado por la empresa.

<sup>24</sup> Ticket Promedio: Término que hace referencia al valor promedio de las transacciones realizadas.



Para cada uno de los conjuntos ya descritos se tiene un grupo de variables que intentan representar su comportamiento, siendo principalmente la frecuencia de incidencia de los eventos uno de los principales elementos a considerar.

#### **4.10.3 Determinación Conjunto de Clientes**

Junto con definir las variables que se utilizarán en el trabajo, se debe determinar cuál será la base de clientes con la que se trabajará. Para realizar esta labor, se trabajó con la base de mora de clientes, esta base fue proporcionada por la Gerencia de Crédito de Ripley y contiene información respecto a los montos y los días de mora de los diferentes clientes de la tienda. Los datos se encuentran detallados a nivel mensual.

Los días de mora son calculados a fin de cada mes, y comprende el número de días que una cuota se mantiene en calidad de impaga posterior a su fecha de vencimiento.

Para efectos de esta memoria, se trabajó el modelo con toda la base de clientes que tiene disponible la empresa, la que contiene cerca de cuatro millones y medio de clientes. Para el desarrollo del modelo, se escogieron 4 meses como base, estos son Junio, Octubre, Noviembre y Diciembre de 2010. A partir de estos meses base, se construyeron las variables resumiendo la información de los 18 meses anteriores al respectivo. Además se construyó la función de desempeño <sup>25</sup> analizando el comportamiento de pago de los 12 meses posteriores al utilizado como *pivote*.

Junto con definir el periodo en mora para clasificar a los clientes se deben explicitar una serie de normas de exclusión con los cuales se filtran a los clientes. Las exclusiones utilizadas en este trabajo fueron las mismas consideradas en el modelo de la empresa y consideran principalmente aspectos como vigencia de la cuenta, defunciones de clientes, antigüedad de la cuenta, entre otros.<sup>26</sup>

#### **4.10.4 Construcción de las Variables**

La construcción de las variables se realizó cruzando diferentes bases de datos dispuestas por la empresa. La construcción se dividió en tres partes, primero se utilizó la base de datos con transacciones históricas de la compañía, desde la cual se construyeron las variables que describen el comportamiento de compra en la tienda a nivel general. Para obtener la información desagregada a nivel de departamento se cruzó la información de transacciones con la base maestra de productos, desde donde

---

<sup>25</sup> Función de Desempeño: Debe entenderse como la clasificación de un cliente en Bueno, Malo o Indeterminado.

<sup>26</sup> Por políticas de confidencialidad se ha omitido el detalle de las exclusiones.

se obtuvo el detalle por área y luego se fue agrupando por cliente y por fecha según el caso.

Posteriormente, para la construcción de las variables que describen el comportamiento en el consumo de servicios financieros, se utilizó una base especial que resume esta información, donde sólo se tuvo que filtrar por periodo de tiempo según el mes correspondiente y luego se agrupó por cada cliente la información.

El mismo procedimiento se siguió con las compras en comercio asociados pero, a diferencia de los anteriores, este proceso resultó ser más engorroso debido a que los códigos de los locales están asociados a una sucursal específica y no necesariamente la cadena completa. Este hecho dificultó la extracción de los datos, ya que para los rubros de comida se debía buscar local por local en una lista con aproximadamente 26 mil locales. Las transacciones se agruparon en 5 grandes áreas de negocios: Bencineras, Comida, Farmacias, Ópticas y Supermercados.

Con la clasificación de clientes y la construcción de las variables completas, es posible realizar el modelamiento y testear las variables construidas.

## **5.- Resultados**

Los resultados obtenidos se presentarán en distintas fases. En primer lugar, se describirán las variables construidas y se presentará el conjunto final de clientes que se obtuvo a partir de la clasificación ya detallada. En segundo lugar, se presenta el conjunto final de variables seleccionadas, el cual considera las 10 variables candidatas a introducirse en el modelo final. Finalmente se presentan los resultados del modelo final que considera la incorporación de las variables seleccionadas.

### **5.1 Base de Clientes**

El trabajo se realizó utilizando dos bases de clientes distintas, una con la clasificación de clientes provista por la empresa y otra con la clasificación desarrollada para efectos de este trabajo. En la sección anterior se detallaron los meses que fueron utilizados como base para construir el desempeño, así como también las principales consideraciones con las cuales se fueron excluyendo clientes de la base final. La Tabla N°7 contiene el número de casos totales con los que se realizó el modelamiento y su correspondiente clasificación en Bueno, Malo e Indeterminado de acuerdo a los criterios desarrollados para este trabajo.

**Tabla N° 7 "Número de Casos Modelo"**

Modelo Prueba		
	Cantidad	Porcentaje
Bueno	2.457.394	77,8%
Indeterminado	363.412	11,5%
Malo	339.315	10,7%
Total	3.160.121	100,0%

**Fuente: Elaboración Propia.**

De los cerca de 18 millones de casos con los que se contaba en un comienzo (considerando el conjunto de los 4 meses utilizados en el estudio), luego de las exclusiones aplicadas, se seleccionaron poco más de tres millones de casos. La cantidad de observaciones obtenidas para las distintas clasificaciones resultó distinta a la utilizada por la empresa<sup>27</sup>, presentando notorias diferencias en la cantidad de Malos, donde la nueva clasificación presenta un número mayor de clientes para esta clasificación. La diferencia encontrada puede ser explicada por el periodo de tiempo considerado para observar el desempeño de pago de los clientes, el criterio "de prueba" observa el comportamiento de pago a un año desde el mes base correspondiente mientras que el criterio de la empresa considera un horizonte de tiempo menor para analizar el desempeño. Dado el menor horizonte de tiempo considerado por el modelo de la compañía, el criterio se vuelve más estricto con el cumplimiento de las condiciones requeridas para ser clasificado como Bueno o Malo. Esta situación provoca que el número de Indeterminados del modelo de la empresa sea mayor, excluyendo el comportamiento de éstos clientes para el modelo.

Otro hecho relevante que incide en las diferencias de ambas clasificaciones es el tratamiento que hace la empresa sobre casos que presentan algunas características especiales sobre los datos observados en el horizonte de tiempo analizado, este tratamiento tiene un impacto en el número de casos que se encuentran en calidad de indeterminados. Aun cuando el número de casos Malos en el criterio utilizado por la empresa es menor, como se verá más adelante, el modelo desarrollado utilizando esta función de desempeño presenta resultados considerablemente mejores que aquel que emplea el "criterio de prueba".

Afinados los criterios de clasificación de los clientes, se procedió a la construcción de las variables sobre las cuales se sostuvo el presente trabajo. En total, se construyeron 157 variables que resumen el comportamiento promedio, observaciones

---

<sup>27</sup> Por políticas de confidencialidad no es adjuntar los resultados de la clasificación de la empresa.

máximas, mínimas, frecuencia de incidencia, la variación en el comportamiento observado, observación inicial y observación final de los clientes para cada una de las categorías de variables ya detalladas. Las variables construidas resumen la información transaccional de los clientes a través de las compras realizadas en los distintos departamentos de la tienda, así como también de la adquisición de algunos servicios financieros ofrecidos por la empresa y otras variables que emanan de la planificación de los pagos que realizan los individuos. Para la selección de las variables se utilizaron diversas fuentes proporcionadas por la empresa, la construcción de las variables fue validada por el equipo de la empresa con el que se realizó el trabajo.

## **5.2 Variables Seleccionadas**

Del total de variables construidas se seleccionó un conjunto de diez variables como candidatas a ser incluidas en el modelo final de la empresa. A continuación, se presentan las aperturas y se grafica el WOE para un grupo de las variables seleccionadas. Para escoger el conjunto final sobre el cual se corrió el modelo, se utilizó el *software* facilitado por la compañía, que como ya se explicó anteriormente permite obtener las aperturas de los intervalos en los cuales se agrupan las variables, el WOE y la valor de la información (VI) de cada una de ellas. Por temas relativos a la confidencialidad de los datos, estos serán presentados con sus nombres originales codificados. De las 157 variables originales finalmente fue seleccionado un conjunto de 10 variables para realizar las pruebas con el modelo y sobre este grupo se seleccionó el conjunto final de variables que se incluyó en el modelo.

El grupo está compuesto por variables provenientes de intercambio en la tienda, la adquisición de algunos servicios financieros, comercio asociado y algunas variables que dan cuenta de la planificación de gastos de los diferentes individuos. Las variables se enlistan de acuerdo a los rangos de las diferentes aperturas y los respectivos gráficos del WOE de cada una.

El listado de variables obtenidas se divide en 3 grupos. El primero, que proviene de la información transaccional en áreas de la tienda y contiene a las variables “CHAR\_216”, “CHAR\_218”, “CHAR\_224” y “CHAR\_228”. El segundo grupo considera variables de comercios asociados e incluye a las variables “CHAR\_220”, “CHAR\_222” y “CHAR\_226”. Finalmente el último grupo se relaciona con la planificación de los pagos de los individuos, donde se encuentran las variables “CHAR\_230”, “CHAR\_232” y “CHAR\_234”.

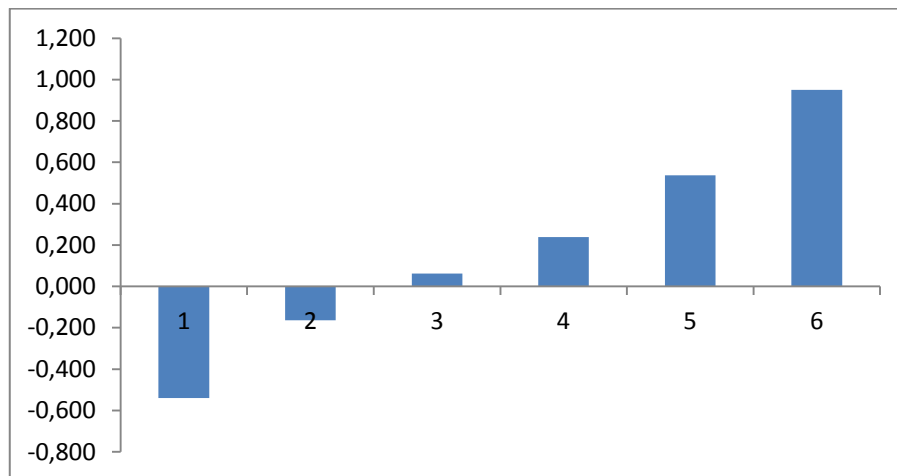
A continuación se enlista parte de las tablas con los intervalos, WoE y la contribución del intervalo al VI final. El resto de las variables se incluye en la tabla de anexos N°2.

**Tabla N° 8 "WOE CHAR\_216"**

CHAR_216	WOE	Contrib. IV
0.0-<77983.0	-0,540	0,075
79653.0-<242410.0	-0,164	0,007
242410.0-<327542.0	0,061	0,000
327542.0-<477921.0	0,239	0,007
477921.0-<915741.0	0,538	0,045
915741.0-<High	0,950	0,062
<b>Total IV</b>	<b>0,197</b>	

Fuente: Elaboración Propia.

**Gráfico N° 1 "WOE CHAR\_216"**



Fuente: Elaboración Propia.

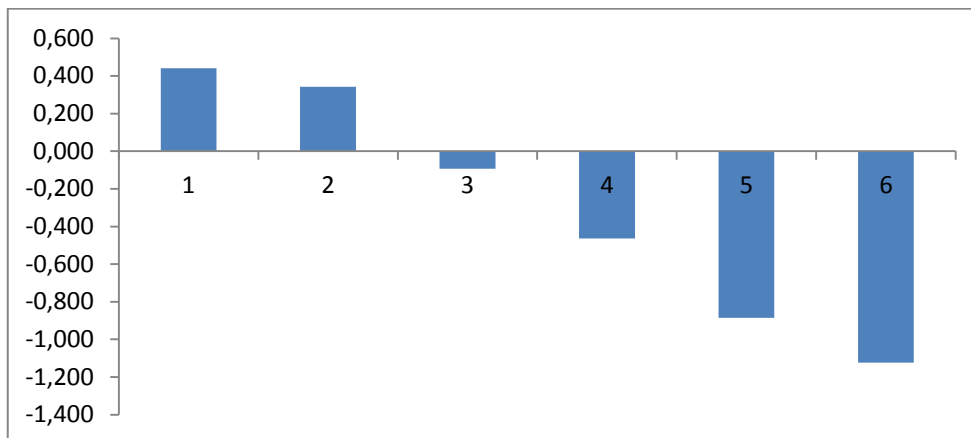
En el gráfico analizado, es posible observar una tendencia lineal de los WOE, esto implica que la relación entre los WOE de cada grupo es significativamente distinta de aquel que corresponde a un grupo siguiente.

**Tabla N° 9 "WOE CHAR\_234"**

CHAR_234	WOE	Contrib. IV
0.0-<2.0	0.441	0.005
2.0-<5.0	0.343	0.043
5.0-<19.0	-0.094	0.004
19.0-<25.0	-0.463	0.019
25.0-<37.0	-0.885	0.043
37.0-<High	-1.124	0.002
<b>Total IV</b>	<b>0,116</b>	

Fuente: Elaboración Propia.

**Gráfico N° 2 "WOE CHAR\_234"**



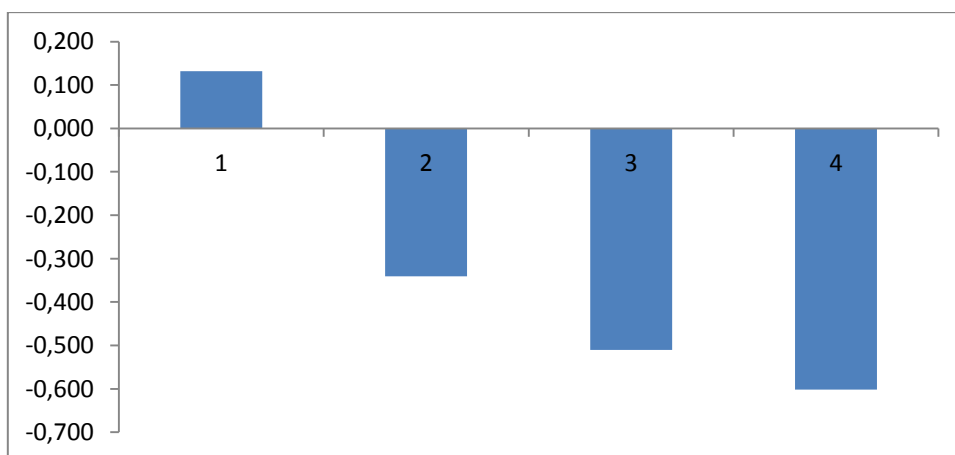
Fuente: Elaboración Propia.

**Tabla N° 10 "WOE CHAR\_224"**

CHAR_224	WOE	Contrib. IV
0.0-<1.0	-0.180	0.024
1.0-<2.0	0.394	0.015
2.0-<3.0	0.657	0.017
3.0-<High	0.818	0.070
<b>Total IV</b>	<b>0,126</b>	

Fuente: Elaboración Propia.

**Gráfico N° 3 "WOE CHAR\_224"**



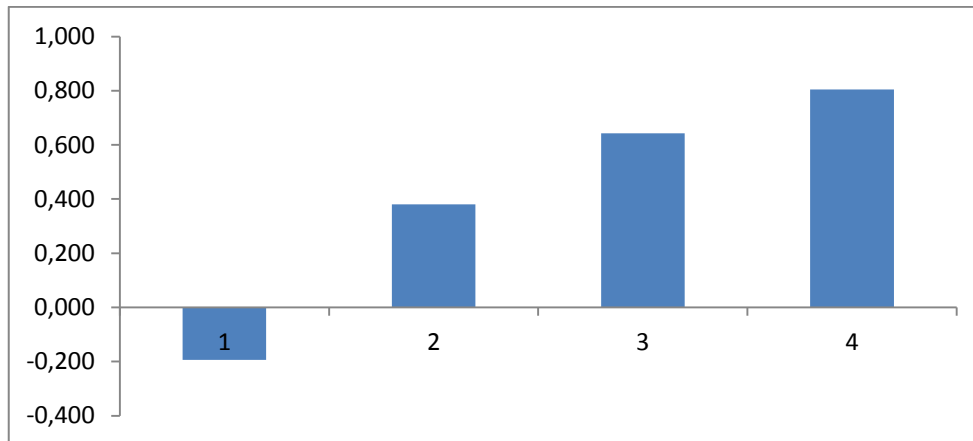
Fuente: Elaboración Propia.

**Tabla N° 11 "WOE CHAR\_222"**

CHAR_222	WOE	Contrib. IV
0.0-<1.0	-0,193	0,028
1.0-<2.0	0,381	0,014
2.0-<3.0	0,643	0,016
3.0-<High	0,804	0,067
<b>Total IV</b>	<b>0,125</b>	

Fuente: Elaboración Propia.

**Gráfico N° 4 "WOE CHAR\_222"**



Fuente: Elaboración Propia.

En general, el conjunto de variables observado presenta buen comportamiento en los WOE observados. Al hablar de un buen comportamiento se hace referencia a que es posible observar una cierta linealidad de los WOE encontrados con respecto al intervalo considerado y estos no se comportan de manera "zigzagueante" mientras se avanzan en los intervalos.

En los gráficos enlistados, es posible observar que los WOE analizados presentan un comportamiento similar entre sí, no encontrándose anomalías en las tendencias que siguen.

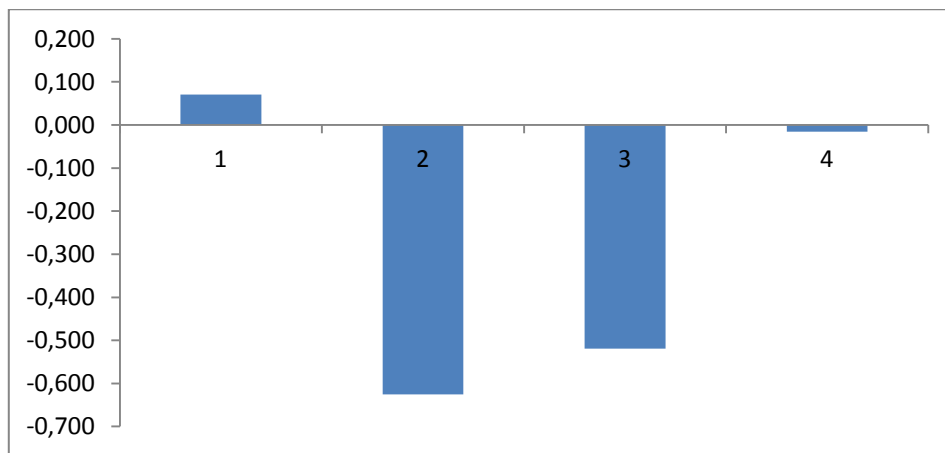
Dentro del conjunto final de variables se seleccionó un par que, si bien es cierto, tenían un buen poder predictivo, al graficar sus WOE se pudo apreciar que este tenía un comportamiento irregular, estas variables son:

**Tabla N° 12 "WOE CHAR\_230"**

CHAR_230	WOE	Contrib. IV
0.0-<1.0	0,070	0,004
1.0-<489920.0	-0,625	0,030
489920.0-<884950.0	-0,519	0,006
884950.0-<High	-0,016	0,000
<b>Total IV</b>	<b>0,041</b>	

Fuente: Elaboración Propia.

**Gráfico N° 5 "WOE CHAR\_230"**



Fuente: Elaboración Propia.

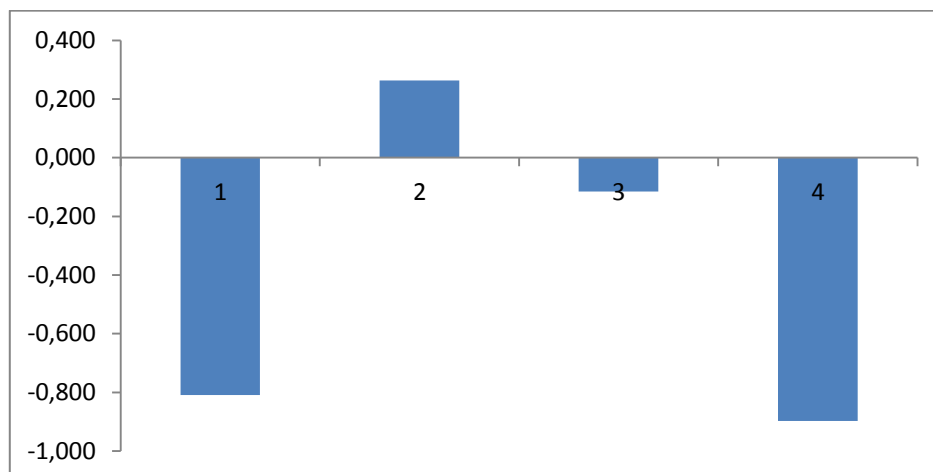
**Tabla N° 13 "WOE CHAR\_232"**

CHAR_232	WOE	Contrib. IV
0.0-<1.0	-0,809	0,070
1.0-<4.0	0,263	0,038
4.0-<25.0	-0,115	0,004
25.0-<High	-0,898	0,020
<b>Total IV</b>	<b>0,131</b>	

Fuente: Elaboración Propia.



**Gráfico N° 6 "WOE CHAR\_232"**



Fuente: Elaboración Propia.

Estas variables presentan un comportamiento anómalo en sus WOE, esto hace que los coeficientes no capturan de buena manera el comportamiento de los clientes, el par de variables que presenta este comportamiento pertenece al último grupo descrito inicialmente. Por esta razón, estas variables no se incluyeron en el conjunto final de variables que se incorporó al modelo de la empresa.

De acuerdo al valor de información que presenta cada una de las variables seleccionadas, estas se ordenan de la siguiente manera:

**Tabla N° 14 "Listado Resumen Variables Seleccionadas"**

Variable	VI
CHAR_216	0,197
CHAR_234	0,133
CHAR_232	0,131
CHAR_222	0,125
CHAR_218	0,086
CHAR_224	0,067
CHAR_226	0,066
CHAR_220	0,057
CHAR_230	0,041
CHAR_228	0,031

Fuente: Elaboración Propia.

Con las variables seleccionadas se procede a realizar la regresión logística para obtener los coeficientes de las variables, calcular el puntaje asociado y con ello medir el poder de discriminación de los modelos construidos.

Un primer análisis sobre las variables de carácter transaccional que se seleccionaron, permite a través de la gráfica de los WOE observar la consistencia de las variables obtenidas. De las diez variables finalmente seleccionadas, ocho presentan consistencia en cuanto al valor calculado de los WOE. Al hablar de consistencia, se hace referencia a la tendencia que se observa respecto de los WOE, donde el ideal es que se pueda apreciar claramente una diferencia entre el WOE un grupo y aquel que corresponde al grupo siguiente. Esta diferencia, junto con ser clara, debe en el ideal lograr que consecutivamente la tendencia del cambio entre un grupo y otro se vaya repitiendo, dicho de otro modo, se deben evitar aquellas variables que presenten un comportamiento “zigzagueante” en los valores obtenidos de los WOE para cada variable.

Cabe señalar que el cálculo del WOE, junto con determinar si una variable es relevante o no, permite posteriormente asignar los puntajes para cada tramo de la variable, por lo que la consistencia del WOE asegura que el puntaje para cada tramo de las variables seleccionadas sea asignado de buena manera.

### **5.3 Resultados Modelamiento**

De acuerdo a la naturaleza binaria de la variable a modelar, para este trabajo se utilizó la regresión logística, cuyas principales características son:

- Asume que la variable objetivo es categórica.
- Puede ser descrita como una función lineal para modelar el “*odd ratio*”.
- Asume la independencia de los errores.
- Se basa en que existe una baja correlación entre los predictores.
- Asume el uso de variables relevantes para modelar el problema.

La fase de modelamiento fue realizado en diferentes etapas; Una primera etapa, orientada a modelar sólo el comportamiento de las variables transaccionales seleccionadas; una segunda etapa orientada a la selección del conjunto final de variables que serán incluidas en el modelo usado por la empresa, esta selección se hace en base al WOE de las variables y al Valor de la Información (VI). Finalmente se corre la regresión logística utilizando las variables finales seleccionadas en la fase anterior. Cabe señalar que tanto para la fase uno como para la última, el modelo fue corrido utilizando la clasificación de clientes empleada por la empresa y la clasificación de clientes desarrollada por este trabajo.

Los resultados de las salidas obtenidas en las diferentes regresiones serán presentados en el mismo orden que se desarrollaron las fases recientemente descritas, siendo el modelo N°1 el que contempla la función de desempeño de la empresa y el modelo N°2 el que utiliza la función de desempeño con los criterios elaborados para este trabajo. En primer lugar se presenta la salida con el listado de los coeficientes

obtenidos, posteriormente la clasificación que entrega la regresión y finalmente el gráfico con la distribución acumulada de los buenos y malos de cada modelo.

La distribución acumulada fue construida a partir del cálculo de los puntajes para cada cliente en función de las características presentadas, este puntaje fue calculado con la fórmula descrita en la Ecuación N° 10. El puntaje por variable se adjunta en la sección de anexos N°3, mientras que el listado con la cantidad de buenos y malos por puntaje total es adjuntado en la sección anexos N°4.

A continuación, se presentan los resultados obtenidos en las diferentes regresiones realizadas.

### 5.3.1 Resultados Modelo de Pruebas N°1

**Tabla N° 15 "Resultados Modelo de Pruebas N°1"**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inferior	Superior
Paso 1 <sup>a</sup>								
WOE_CHAR_216	-,634	,007	7380,698	1	,000	,531	,523	,538
WOE_CHAR_218	-,376	,011	1117,770	1	,000	,686	,672	,702
WOE_CHAR_220	-,328	,014	525,223	1	,000	,720	,700	,741
WOE_CHAR_222	-,532	,010	3109,580	1	,000	,587	,576	,598
WOE_CHAR_224	-,602	,068	77,997	1	,000	,548	,480	,626
WOE_CHAR_226	-,680	,069	95,703	1	,000	,507	,442	,581
WOE_CHAR_228	-,995	,020	2532,938	1	,000	,370	,356	,384
WOE_CHAR_230	-,127	,026	23,476	1	,000	,881	,837	,927
WOE_CHAR_232	-,399	,012	1176,582	1	,000	,671	,656	,687
WOE_CHAR_234	-,412	,012	1154,000	1	,000	,662	,647	,678
Constante	-2,789	,003	866631,709	1	,000	,061		

a. Variable(s) introducida(s) en el paso 1: WOE\_CHAR\_216, WOE\_CHAR\_218, WOE\_CHAR\_220, WOE\_CHAR\_222, WOE\_CHAR\_224, WOE\_CHAR\_226, WOE\_CHAR\_228, WOE\_CHAR\_230, WOE\_CHAR\_232, WOE\_CHAR\_234.

**Fuente: Elaboración Propia.**

**Tabla N° 16 "Resultados Clasificación Modelo de Pruebas N°1"**

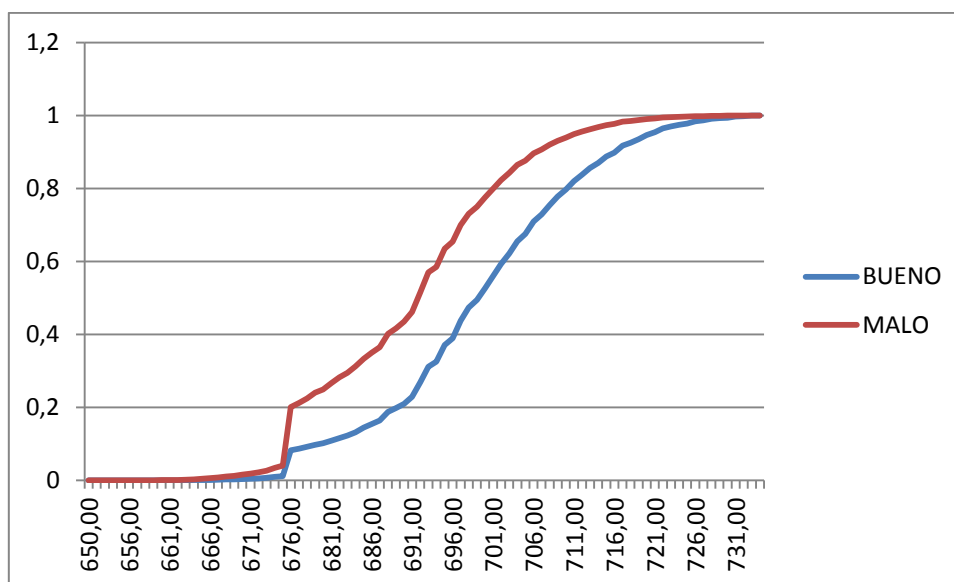
**Tabla de clasificación<sup>a</sup>**

Observado			Pronosticado		Porcentaje correcto
			BUENO_MALO		
			,00	1,00	
Paso 1	BUENO_MALO	,00	90080	47539	65,5
		1,00	55182	82437	59,9
Porcentaje global					62,7

a. El valor de corte es ,500

Fuente: Elaboración Propia.

**Gráfico N° 7 "Distribución Acumulada Modelo de Pruebas N°1"**



Fuente: Elaboración Propia.

El modelo de prueba N°1 incluye el conjunto de variables transaccionales y la clasificación de clientes utilizada por la empresa. El Indicador K-S de este modelo es 24,42%. El valor de este indicador implica que el modelo en general no es bueno lo que habla del poco poder predictivo de las variables transaccionales por sí solas.

### 5.3.2 Resultados Modelo de Prueba N°2

El modelo de pruebas N°2 se realiza utilizando la clasificación de clientes que se elaboró para este trabajo el que, como ya se explicó en el capítulo anterior, difiere del utilizado por la empresa en el número de clientes que considera.

**Tabla N° 17 "Resultado Modelo de Pruebas N°2"**

		Variables en la ecuación						I.C. 95% para EXP(B)	
		B	E.T.	Wald	gl	Sig.	Exp(B)	Inferior	Superior
Paso 1 <sup>a</sup>	WOE_CHAR_216	-1,720	,027	3968,536	1	,000	,179	,294	,305
	WOE_CHAR_218	-,790	,009	7398,459	1	,000	,454	,617	,644
	WOE_CHAR_220	-,137	,062	4,910	1	,027	,872	,627	,660
	WOE_CHAR_222	-,699	,015	2154,142	1	,000	,497	,564	,583
	WOE_CHAR_224	-,499	,013	1476,846	1	,000	,607	,518	,610
	WOE_CHAR_226	-,637	,016	1509,410	1	,000	,529	,729	,860
	WOE_CHAR_228	-,708	,013	3034,079	1	,000	,493	,489	,508
	WOE_CHAR_230	-,572	,009	3671,551	1	,000	,564	,529	,547
	WOE_CHAR_232	-,692	,062	126,400	1	,000	,501	,624	,640
	WOE_CHAR_234	-,663	,011	3852,139	1	,000	,515	,402	,413
	Constante	-,144	,003	1919,370	1	,000	,866		

a. Variable(s) introducida(s) en el paso 1: WOE\_CHAR\_216, WOE\_CHAR\_218, WOE\_CHAR\_220, WOE\_CHAR\_222, WOE\_CHAR\_224, WOE\_CHAR\_226, WOE\_CHAR\_228, WOE\_CHAR\_230, WOE\_CHAR\_232, WOE\_CHAR\_234.

Fuente: Elaboración Propia.

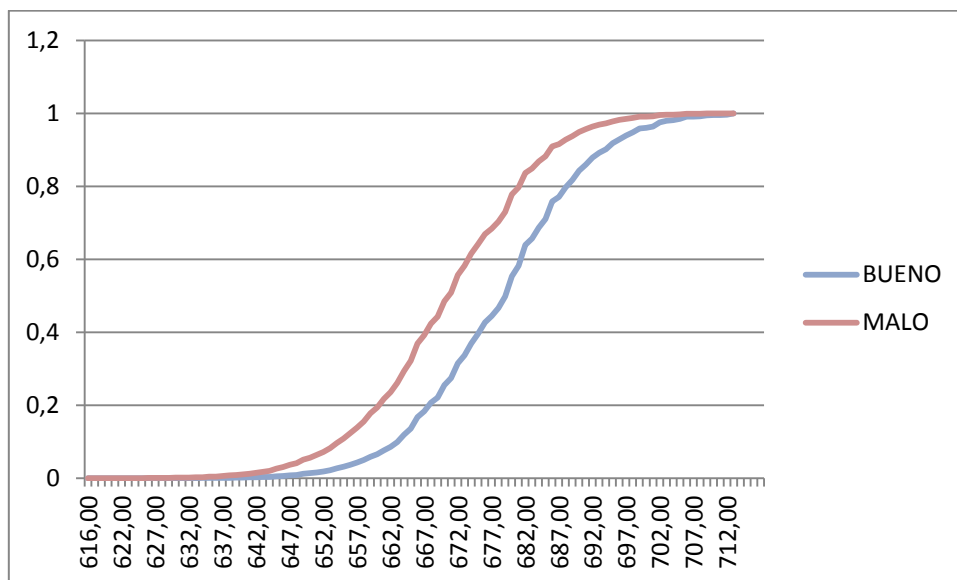
**Tabla N° 18 "Resultados Clasificación Modelo de Prueba N°2"**

		Tabla de clasificación <sup>a</sup>		
		Pronosticado		
		BUENO_MALO		Porcentaje correcto
Observado		,00	1,00	
Paso 1	BUENO_MALO ,00	219850	119465	64,8
	1,00	139679	199636	58,8
Porcentaje global				61,8

a. El valor de corte es ,500

Fuente: Elaboración Propia.

**Gráfico N° 8 "Distribución Acumulada Modelo de Pruebas N°2"**



Fuente: Elaboración Propia.

El resultado del indicador K-S para este modelo es de 22,35%. Al igual que en el modelo anterior, no se obtuvo un buen indicador. Dado que los conjuntos de clientes presentan diferencias entre sí, los resultados obtenidos no son directamente comparables. Sin embargo, la similitud en los resultados obtenidos indican que las variables transaccionales por sí solas no son buenos estimadores para un modelo de comportamiento como el que se está estudiando.

Los dos primeros modelos fueron utilizados para entender cómo se comportaban este tipo de variables dentro de un modelo de estas características y para entender cuáles de éstas tienen mayor relevancia. De las diez variables preliminares, se escogieron cuatro para ser incorporadas dentro del modelo de riesgo que utiliza la compañía. Las variables escogidas se seleccionaron en base al Valor de la Información (VI) que presentaban (definido en el capítulo N°3), al comportamiento observado en el modelamiento preliminar y a la consistencia observada en la gráfica de los WOE de la sección anterior. Dentro de este grupo es posible encontrar variables que describen parte del comportamiento agregado de compras en la tienda, compras en un departamento específico, cierto comportamiento en un comercio asociado y una variable que considera la planificación futura de los pagos.

El listado final con las variables seleccionadas es el siguiente:

**Tabla N° 19 "Listado Final Variables"**

<b>Listado Final</b>
WOE_CHAR_200
WOE_CHAR_202
WOE_CHAR_204
WOE_CHAR_208
WOE_CHAR_210
WOE_CHAR_212
WOE_CHAR_214
WOE_CHAR_216
WOE_CHAR_222
WOE_CHAR_224
WOE_CHAR_234

**Fuente: Elaboración Propia.**

El listado final contiene once variables, de las cuales las siete primeras corresponden al modelo ya utilizado por la empresa y las cuatro finales son de carácter transaccional.

#### **5.3.4 Resultados Modelo Mixto N°1**

El modelo Mixto N°1 contiene siete de las ocho variables transaccionales utilizadas por la empresa y cuatro de las diez variables finales de carácter transaccional. Utiliza la clasificación de clientes de la empresa. Las comparaciones fueron realizadas en base al conjunto que incluye sólo a las variables de la empresa a fin de medir el impacto real que tiene la adición de variables transaccionales en el modelo. Los resultados obtenidos se adjuntan a continuación:

**Tabla N° 20 "Resultados Modelo Mixto N°1"**

		Variables en la ecuación							I.C. 95% para EXP(B)	
		B	E.T.	Wald	gl	Sig.	Exp(B)	Inferior	Superior	
Paso 1 <sup>a</sup>	WOE_CHAR_200	-,823	,008	10756,586	1	,000	,439	,462	,471	
	WOE_CHAR_202	-,675	,010	4389,663	1	,000	,509	,538	,552	
	WOE_CHAR_204	-,747	,007	11294,605	1	,000	,474	,518	,526	
	WOE_CHAR_208	-,662	,009	4940,770	1	,000	,516	,544	,557	
	WOE_CHAR_210	-,392	,010	1568,581	1	,000	,675	,697	,713	
	WOE_CHAR_212	-,307	,010	1023,911	1	,000	,736	,773	,791	
	WOE_CHAR_214	-,222	,010	521,445	1	,000	,801	,631	,648	
	WOE_CHAR_216	-,883	,023	1455,796	1	,000	,414	,586	,604	
	WOE_CHAR_222	-1,027	,027	1455,400	1	,000	,358	,664	,692	
	WOE_CHAR_224	-,472	,014	1069,606	1	,000	,624	,361	,385	
	WOE_CHAR_234	-,362	,016	512,858	1	,000	,696	,631	,656	
	Constante	-,096	,007	210,631	1	,000	,908			

a. Variable(s) introducida(s) en el paso 1: WOE\_CHAR\_200, WOE\_CHAR\_202, WOE\_CHAR\_204, WOE\_CHAR\_208, WOE\_CHAR\_210, WOE\_CHAR\_212, WOE\_CHAR\_214, WOE\_CHAR\_216, WOE\_CHAR\_222, WOE\_CHAR\_224, WOE\_CHAR\_234.

Fuente: Elaboración Propia.

**Tabla N° 21 "Resultados Clasificación Modelo Híbrido N°1"**

Tabla de clasificación<sup>a</sup>

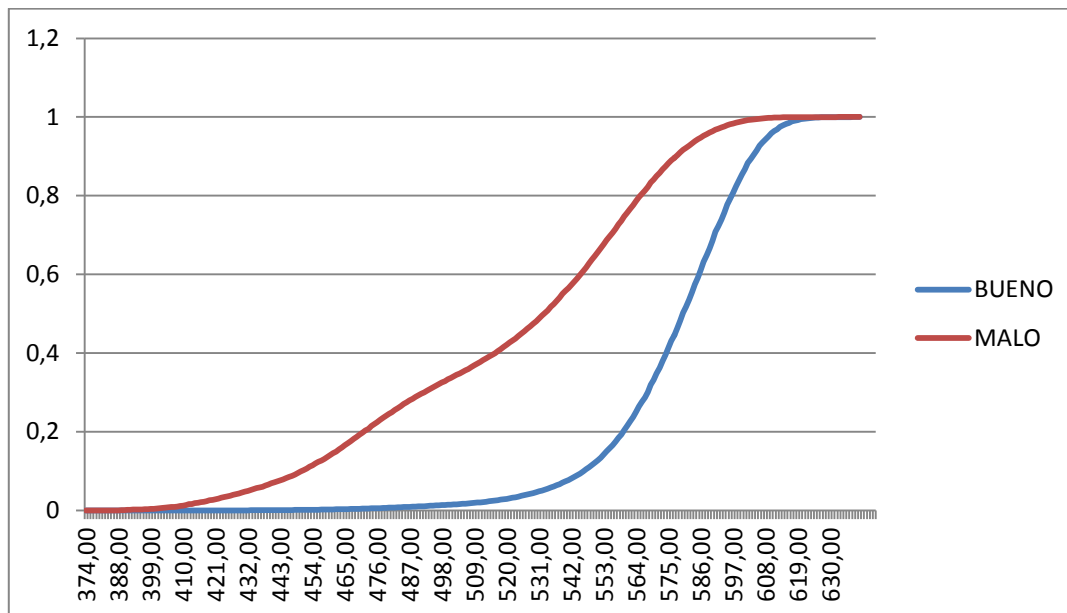
Observado			Pronosticado		
			BUENO_MALO		Porcentaje correcto
,00	1,00				
Paso 1	BUENO_MALO	,00	113878	23741	82,7
		1,00	39051	98568	71,6
Porcentaje global					77,2

a. El valor de corte es ,500

Fuente: Elaboración Propia.



**Gráfico N° 9 "Distribución Acumulada Modelo Mixto N°1"**



Fuente: Elaboración Propia.

El indicador K-S obtenido para este modelo fue de un 54,33%. De la gráfica anterior, es posible apreciar que la cantidad de clientes Buenos crece significativamente a partir de los 530 puntos.

### 5.3.5 Resultados Modelo Mixto N°2

Este modelo utiliza las mismas variables que el modelo anterior, pero su diferencia radica en que es utilizada la clasificación de clientes construida especialmente para este trabajo, los resultados se adjuntan a continuación:

**Tabla N° 22 "Resultados Modelo Mixto N°2"**

		Variables en la ecuación						I.C. 95% para EXP(B)	
		B	E.T.	Wald	gl	Sig.	Exp(B)	Inferior	Superior
Paso 1 <sup>a</sup>	WOE_CHAR_200	-,925	,007	15623,428	1	,000	,397	,421	,430
	WOE_CHAR_202	-,361	,005	5173,275	1	,000	,697	,512	,521
	WOE_CHAR_204	-,120	,007	308,178	1	,000	,887	,469	,475
	WOE_CHAR_208	-,590	,006	11257,363	1	,000	,554	,561	,570
	WOE_CHAR_210	-,469	,006	6108,826	1	,000	,626	,788	,801
	WOE_CHAR_212	-,405	,007	3323,581	1	,000	,667	,724	,737
	WOE_CHAR_214	-,793	,007	11444,658	1	,000	,453	,855	,868
	WOE_CHAR_216	-,1611	,029	3031,095	1	,000	,200	,421	,439
	WOE_CHAR_222	-,764	,010	6359,048	1	,000	,466	,634	,656
	WOE_CHAR_224	-,695	,011	4219,467	1	,000	,499	,494	,509
	WOE_CHAR_234	-,546	,011	2277,668	1	,000	,579	,509	,523
	Constante	-,110	,007	257,126	1	,000	,896		

a. Variable(s) introducida(s) en el paso 1: WOE\_CHAR\_200, WOE\_CHAR\_202, WOE\_CHAR\_204, WOE\_CHAR\_208, WOE\_CHAR\_210, WOE\_CHAR\_212, WOE\_CHAR\_214, WOE\_CHAR\_216, WOE\_CHAR\_222, WOE\_CHAR\_224, WOE\_CHAR\_234.

Fuente: Elaboración Propia.

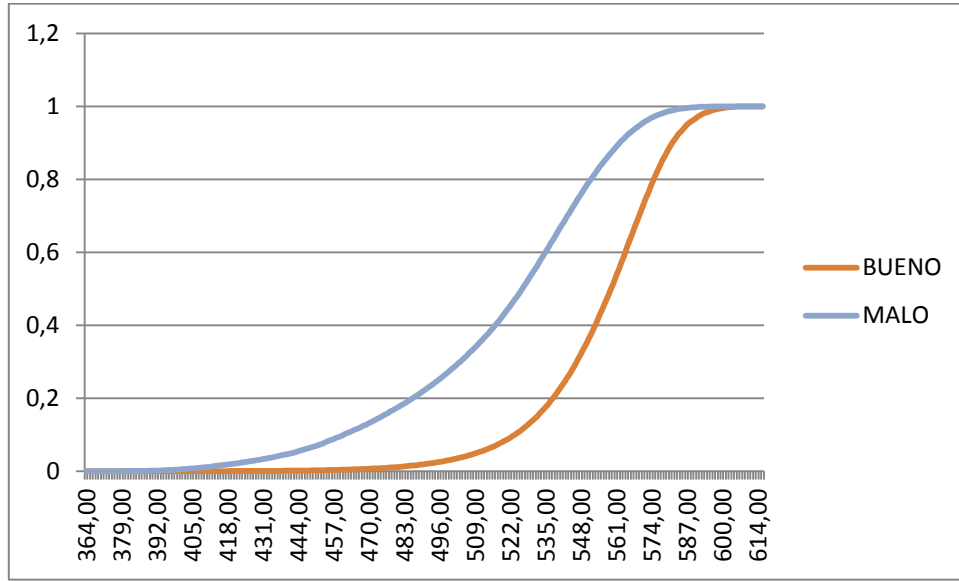
**Tabla N° 23 "Resultados Clasificación Modelo Mixto N°2"**

		Tabla de clasificación <sup>a</sup>			
		Pronosticado		Porcentaje correcto	
Observado		BUENO_MALO	1,00		
Paso 1	BUENO_MALO	,00	1,00		
		256278	83037	75,5	
		124395	214920	63,3	
	Porcentaje global			69,4	

a. El valor de corte es ,500

Fuente: Elaboración Propia.

**Gráfico N° 10 "Distribución Acumulada Modelo Mixto N°2"**



Fuente: Elaboración Propia.

El indicador K-S obtenido para este modelo fue de un 38,8%. De acuerdo a la gráfica de la distribución acumulada se aprecia una clara distinción entre las curvas de cada una de las clasificaciones de clientes.

#### 5.4- Análisis de Resultados

Como se detalló en la sección anterior de este trabajo, las variables seleccionadas fueron construidas resumiendo la información transaccional de los clientes que utilizaron como medio de pago la tarjeta de crédito de la tienda en diversos departamentos de la misma y en comercios asociados. Lo que se buscó con este trabajo es conocer si alguna variable de estas características ayuda a mejorar el poder predictivo del modelo que en la actualidad utiliza la empresa.

Para lograr este objetivo se consideraron desde una perspectiva agregada las transacciones de los clientes, para luego desagregar la información a nivel de área dentro de la tienda. Por otro lado, se exploró la información de los servicios financieros que ofrece la empresa, así como también la información que proviene de las ventas en otro tipo de comercio, paralelamente se analizó el comportamiento referido a la planificación de pagos que realizan los clientes en sus respectivas transacciones.

Esta heterogeneidad de fuentes de información y de tipos de variables que se construyeron, se vio reflejada en el conjunto final de variables seleccionadas, donde cada conjunto descrito aportó con variables al grupo que finalmente se seleccionó. El conjunto final está constituido por variables relacionadas con la frecuencia de visita a ciertos departamentos de la tienda, así como también la frecuencia de visita a cierto tipo de comercio con rubro similar al de la empresa, se encuentran también algunas

variables que dan cuenta de la planificación de los distintos clientes en los pagos que realizarán, otras pertenecientes al negocio financiero de la empresa y finalmente un grupo que considera cantidades a nivel agregado.

Desde la perspectiva del poder predictivo que tienen las variables, el cual se determina por el “valor de la información” (VI) calculado a partir de los WOE, es posible apreciar que el VI de la mayoría de las variables obtenidas indican que éstas poseen un poder predictivo débil<sup>28</sup>. Sin embargo, las cuatro variables que se incluyeron tienen un VI que permite pensar que pueden ser relevantes al incorporarlas dentro del modelo ya utilizado.

El modelamiento fue realizado utilizando como regresores el WOE de las variables construidas. Se modeló de esta forma, puesto que así se aplacan problemas que emanan de la naturaleza de los datos, además esto permite capturar de mejor manera la tendencia que tienen las variables. El supuesto sobre el que descansa esta forma de modelar es el hecho de que la agrupación de los tramos para cada una de las variables fue realizada de manera correcta.

Una vez corridos los modelos utilizados para la estimación de los coeficientes es posible observar que para las cuatro regresiones realizadas, los coeficientes estimados para cada una de las variables explicativas resultaron ser estadísticamente significativos contra la hipótesis nula de que los parámetros sean iguales a cero. Analizando los modelos de prueba que se utilizaron para medir cuán explicativas son las variables transaccionales contra la variable objetivo, es posible observar en las tablas adjuntas en la sección N°6 de los anexos, el modelo que utiliza la función de desempeño empleada por la empresa tiene mejores valores en las pruebas de bondad que se realizaron, esto también se ve reflejado en el indicador K-S obtenido para cada modelo.

---

<sup>28</sup> Ver Tabla N°14 “Listado Resumen Variables Seleccionadas”

**Tabla N° 24 "Indicador K-S Modelo de Prueba"**

Modelo	Indicador K-S
Modelo de Pruebas N°1	24,42%
Modelo de Pruebas N°2	22,35%

**Fuente: Elaboración Propia.**

Como se aprecia en la Tabla N°24, el K-S obtenido en el modelo de pruebas N°1, que utiliza la función de desempeño de la empresa tiene un mejor poder de discriminación con respecto al modelo que utiliza la otra clasificación.

Sin embargo, ambos modelos presentan un poder predictivo bajo en comparación a otros que utilizan variables de comportamiento como predictores del modelo. El resultado obtenido tiene lógica tomando en cuenta que utiliza variables que contemplan transacciones para explicar un fenómeno asociado al comportamiento de pago de los individuos de la muestra. El motivo por el cual se realizó este modelamiento de prueba es de carácter exploratorio y busca analizar cómo se comportan las variables transaccionales por sí solas en un modelo de estas características, además otorgó una mirada complementaria para seleccionar las variables que finalmente se incorporaron en el modelo.

Analizando la matriz de correlaciones adjuntada en la sección N°5 de anexos, se debe notar que las variables incluidas al modelo tienen una baja correlación entre sí y con las del modelo de la empresa siendo la mayor la encontrada entre la variable CHAR\_234 Y CHAR\_202 que es -0.336, ambas variables están relacionadas con características de la planificación de los pagos de los clientes. También se observa una alta correlación entre variables del modelo utilizado por la empresa, siendo ésta de -0.728 para un conjunto de variables relacionadas con el nivel de deuda de la empresa.

Al igual que en los resultados obtenidos para los modelos de prueba, los indicadores de bondad del "modelo mixto" N°1 (con la clasificación de clientes de la compañía), resultan ser mejores que los del modelo N°2, lo mismo se aplica para el indicador K-S obtenido.

**Tabla N° 25 "Indicador K-S Modelo Mixto"**

<b>Modelo</b>	<b>Indicador K-S</b>
Modelo Mixto N°1	54,33%
Modelo de Mixto N°2	38,80%

**Fuente: Elaboración Propia.**

Las diferencias obtenidas en el indicador son mayores que para el modelo anterior. Este resultado le permite de cierta manera a la compañía validar su función de desempeño frente a otra de carácter más "teórico", pero plantea la duda sobre si los resultados obtenidos se deben efectivamente a que el modelo captura de buena manera el comportamiento de aquellos individuos clasificados como malos, o ve influenciado su comportamiento por el bajo número de casos malos que presenta esta clasificación.

El objetivo de este estudio era determinar si las variables transaccionales tenían impacto en un modelo crediticio y permiten complementar las variables de comportamiento utilizadas mejorando el poder predictivo de este.

El K-S obtenido en el "Modelo Mixto N°1" de 54,33% presenta dominancia sobre los resultados del modelo de la empresa, obteniéndose un mejor resultado que el que se obtiene con el modelo de la empresa. Este resultado confirma la hipótesis inicial de esta memoria, donde se plantea que la incorporación de variables transaccionales mejora el poder predictivo del modelo utilizado por la empresa,

Ora medida de la sensibilidad de un modelo se basa en el área bajo la Curva (AUC) de un modelo, que mide la sensibilidad de la representación de un modelo de un modelo para una clasificación binaria, el puntaje del indicador va desde 0,5 (modelo malo) hasta 1 (modelo muy bueno). El valor obtenido para cada modelo se adjunta en la siguiente tabla:

**Tabla N° 26 "Indicador AUC"**

<b>Tipo de Modelo</b>	<b>AUC</b>
<b>Modelo Prueba N°1</b>	0,65
<b>Modelo Prueba N°2</b>	0,66
<b>Modelo Mixto N°1</b>	0,85
<b>Modelo Mixto N°2</b>	0,76

**Fuente: Elaboración Propia.**

De la tabla N°26, es posible apreciar que al igual que el resultado obtenido para los indicadores K-S en los modelos de prueba, los indicadores AUC de estos modelos

se encuentran en un intervalo que se puede clasificar como regular, lo que habla de la poco poder predictivo que tienen las variables de carácter transaccionales por sí solas. Por otro lado, al incorporar el conjunto de variables transaccionales dentro del modelo utilizado por la compañía, se aprecia una mejora de los resultados y el valor de los indicadores se considera bueno. Al igual que en el indicador anterior, los resultados obtenidos, indican que el modelo utilizado por la compañía discrimina mejor que el modelo propuesto, pero las comparaciones no pueden ser realizadas de manera tan directa puesto que el número de clientes en ambos modelos difieren entre sí.

Sobre el conjunto final seleccionado, las cuatro variables tienen diferentes grupos de procedencia, encontrándose en ellas una que proviene del historial mensual de compras, otra proveniente de una sección de la empresa, una que incorpora información sobre la planificación de los pagos de deuda y otra que describe el historial en un rubro similar al de la empresa.

Analizando los puntajes calculados para el último modelo, es posible inferir que para las dos primeras variables seleccionadas, el modelo “premia” los tramos superiores del comportamiento observado, mientras que para el par final, se produce el efecto contrario y se “premia” a los tramos inferiores.

## **6.- Conclusiones y Recomendaciones**

En la presente memoria se pudo estudiar el modelo de riesgo crediticio utilizado por una conocida empresa del retail. Este modelo utiliza un conjunto acotado de variables que describen distintos aspectos del comportamiento de pago de los clientes. El contexto en el que se desenvuelve la empresa, obliga a estar constantemente revaluando los modelos utilizados, de tal manera de hacer frente a los vertiginosos cambios en los patrones de consumo que presentan los clientes. Es por este motivo, que el trabajo realizado presenta el valor de analizar un conjunto de variables que en la actualidad no son consideradas por el actual modelo.

El estudio fue llevado a cabo utilizando una serie de *software* y de herramientas estadísticas que aseguran la calidad de los resultados obtenidos. La metodología empleada resume gran parte de los estudios que se han llevado a cabo en estas materias y su utilización permitió la obtención de resultados relevantes de acuerdo a la naturaleza del fenómeno estudiado.

El trabajo desarrollado condensó la información disponible dentro de la empresa en un conjunto de variables que sirvieron para modelar el fenómeno del comportamiento de pago que presentan los clientes, en este sentido se concluye que el objetivo de lograr la construcción de una base adecuada para realizar el modelamiento fue cumplido.

En base a los resultados obtenidos, es posible señalar que se comprueba la hipótesis planteada de que las variables transaccionales tienen impacto en el modelo crediticio de la compañía y mejora el poder predictivo del mismo. Al inicio de este trabajo, se planteó como objetivo lograr que la incorporación de este tipo de variables permitiera una mejora de un 10 % en el indicador K-S del modelo, sin embargo esto no pudo ser logrado y sólo se logró una mejora menor a la esperada. Lejos de ser un resultado poco satisfactorio, el hecho de comprobar la existencia de un impacto abre la posibilidad de seguir investigando esta línea de trabajo, no descartando que al controlar este tipo de variables por otras como el cupo o el ingreso se obtengan mejores resultados.

Los resultados obtenidos permitieron también a la empresa validar la función de desempeño utilizada por ellos frente a una clasificación realizada en base a una definición tradicional utilizada en la literatura consultada. Esto se basa en que para los diferentes modelos utilizados, aquel que consideraba la función de desempeño de la empresa obtuvo mejores resultados de acuerdo a los diferentes indicadores de bondad del que entrega el modelo.

En total se corrieron cuatro modelos logísticos, para cada uno se estimaron los coeficientes de los distintos regresores considerados, siendo estos estadísticamente significativos en todos los modelos. Los cálculos derivados de la estimación de los parámetros permitieron la elaboración de un puntaje para cada tramo de las diferentes variables obtenidas. Estos puntajes son consistentes con lo que uno esperaría para cada una de las variables, además de estar alineados con los parámetros utilizados por la compañía, por lo que se concluye que los modelos de prueba que se plantearon construir en un comienzo fueron realizados de buena manera.

La riqueza de este trabajo radica en el hecho de haber analizado variables que incorporan información proveniente de diversas áreas de la empresa y que describen parte del comportamiento transaccional que tienen los clientes dentro de la tienda. Es por esto que el resultado obtenido, desde la perspectiva de las variables seleccionadas, es relevante para futuros análisis respecto de fuentes que permitan enriquecer el modelo más allá de las tradicionalmente utilizadas. Por otro lado entrega un análisis del modelo con una nueva clasificación de clientes, lo que le permite a la empresa comparar sus resultados frente a una nueva clasificación y analizar posibilidades de mejora para el modelo.

Cabe señalar, que los modelos desarrollados en este trabajo tuvieron un carácter meramente exploratorio, por lo que el conjunto final de variables obtenidas entrega una serie de lineamientos para futuras investigaciones que profundicen el conocimiento sobre el comportamiento de estas variables y sobre la aplicación efectiva en un modelo de riesgo crediticio



## 7.- Bibliografía y Fuentes de Información

- ANDERSON, Raymond. *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk, Management and Decision Automation*. Oxford: Oxford University Press, 2007. ISBN 978-0-19-922640-5
- BANCO CENTRAL DE CHILE, 2007 Capítulo III.J Compendio de Normas Financiera [en línea] Santiago, Chile, < <http://www.bcentral.cl/normativa/normas-financieras/pdf/CapIIIJ1.PDF> > [Consulta 12 mayo 2012]
- HULL, John C. *Options, Futures and Others Derivatives* 5ta Edition EE.UU. Prentice Hall, 2003 ISBN-10: 0130090565
- HUYEN, Dinh T., KLEIMEIER, Stefanie. *Credit Scoring for Vietnam's Retail Banking Market: Implementation and Implications for Transactional versus Relationship Lending*. Maastricht, Maastricht University. 2006.
- KHANDANI, Amir E., ADLAR, Kim J., LO. Andrew W. *Consumer Credit Risk Models: A Machine-Learning Algorithm*. Cambridge, Massachusetts Institute of Technology. 2010.
- SIDDIQI, Naeem. *Designing Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring*. New Jersey: John Wiley & Sons, Inc., 2006. ISBN-13: 978-0-471-75451-0
- SUPERINTENDECIA DE BANCOS E INSTITUCIONES FINANCIERAS, 2011, Emisores de Deuda no Bancarios Informe trimestral de tarjetas de créditos no Bancarias. [en línea] Santiago, Chile, < <http://www.sbif.cl/sbifweb/servlet/InfoFinanciera?indice=4.1&idCategoria=2129&ti pocont=0> > [Consulta 12 mayo 2012]
- THOMAS, Lyn. *Consumer Credit Models: Pricing, Profit and Portfolios*. Oxford: Oxford University Press, 2009. ISBN 978-0-19-923213-0
- THOMAS, Lyn. A survey of credit and Behavioral scoring: forecasting financial risk of lending to consumers. *International Journal of Forecasting* (16):149-172 Edinburgh: University of Edinburgh.
- VISSING-JORGENSEN, Annette *Consumer Credit: Learning Your Customer's Default Risk from What (S) he buys*. Chicago, Illinois Kellogg School of Management. 2011.

## **Anexos N°1 “Obligaciones Emisor-Titular Tarjetas de Crédito”**

### **Contratos del emisor con el titular de la tarjeta de crédito.** <sup>29</sup>

Los emisores de tarjetas de crédito deberán suscribir o renovar con cada titular de dicho instrumento, un “Contrato de afiliación al sistema y uso de la tarjeta”, que deberá contemplar los siguientes contenidos mínimos:

- 1) el plazo o condiciones de vigencia del contrato;
- 2) el límite de crédito autorizado por el período contratado. Las modificaciones a ese límite deberán ser informadas por escrito al titular y, si estas consisten en una disminución del cupo pactado, en la notificación correspondiente deberán indicarse las causas objetivas en que se funda esa determinación, las cuales deberán estar previamente pactadas con el tarjetahabiente en el respectivo contrato;
- 3) la fecha de emisión de estados de cuenta y de vencimiento de la respectiva obligación de pago del titular o usuario;
- 4) las modalidades y condiciones aplicables al cobro de comisiones y/o cargos e intereses, las que podrán ser modificadas por el emisor previo aviso al titular en el estado de cuenta y en las pizarras informativas en locales del emisor;
- 5) el costo de comisiones y/o cargos por mantención de la Tarjeta, las que podrán ser modificadas previo aviso del emisor al titular;
- 6) las medidas de seguridad relacionadas con el uso de la Tarjeta y los procedimientos y responsabilidades en caso de robo, hurto, pérdida, adulteración o falsificación de la misma;
- 7) la resolución de controversias;
- 8) los requisitos y condiciones aplicables respecto del término del contrato, incluidas las causales de término unilateral del mismo; y,
- 9) los derechos conferidos al titular o usuario de que trata el párrafo 4° de la Ley 19.496, en materia de normas de equidad en las estipulaciones y en el cumplimiento de los contratos de adhesión

---

<sup>29</sup> Circular N°17 Superintendencia de Bancos e Instituciones Financieras.

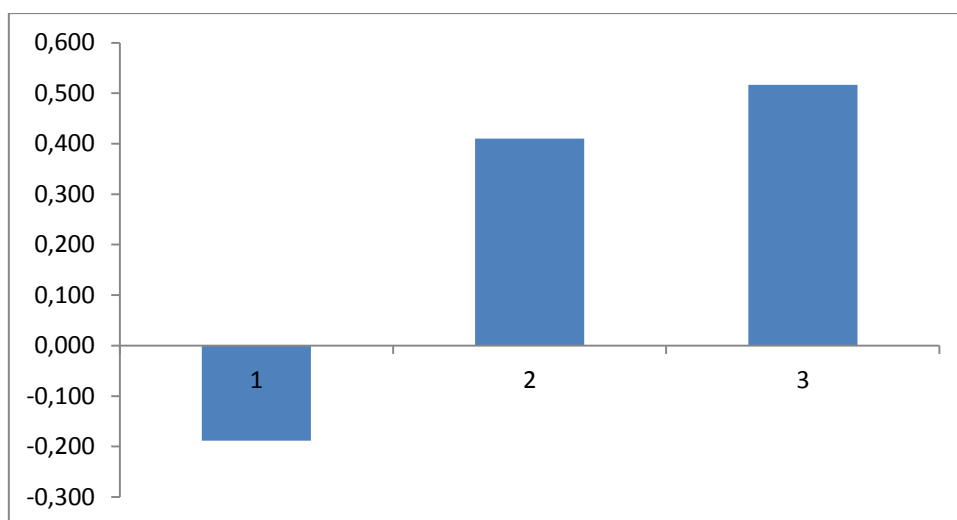
## Anexos N°2 "WOE Variables de Muestra"

### Tabla N° 27 "WOE CHAR\_218"

CHAR_218	WOE	Contrib. IV
0.0-<1.0	-0,188	0,025
1.0-<3.0	0,410	0,027
3.0-<High	0,517	0,034
<b>Total IV</b>	<b>0,086</b>	

Fuente: Elaboración Propia.

### Gráfico N° 11 "WOE CHAR\_218"



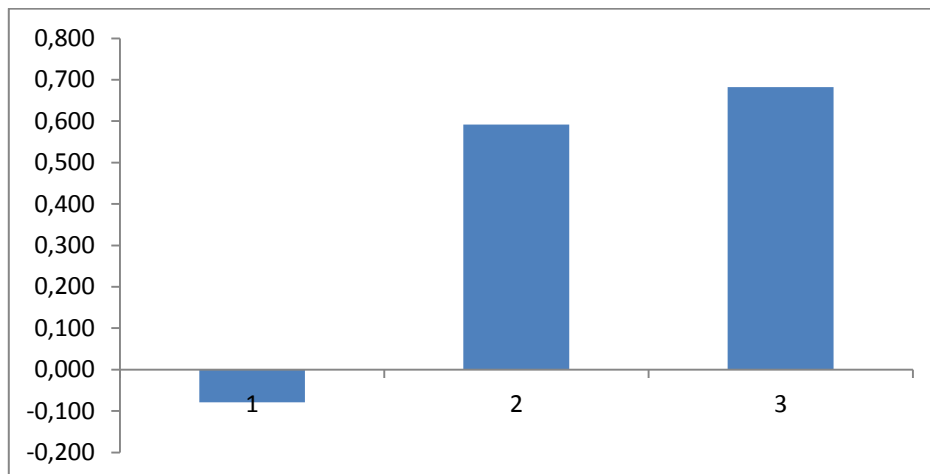
Fuente: Elaboración Propia.

### Tabla N° 28 "WOE CHAR\_228"

CHAR_228	WOE	Contrib. IV
0.0-<1.0	-0,079	0,005
1.0-<2.0	0,592	0,025
2.0-<High	0,682	0,027
<b>Total IV</b>	<b>0,058</b>	

Fuente: Elaboración Propia.

**Gráfico N° 12 "WOE CHAR\_228"**



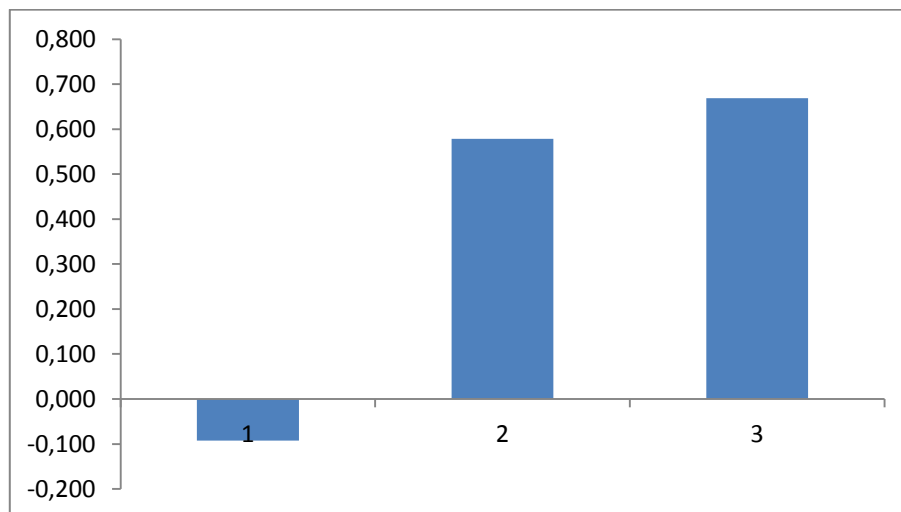
Fuente: Elaboración Propia.

**Tabla N° 29 "WOE CHAR\_220"**

CHAR_220	WOE	Contrib. IV
0.0-<1.0	-0,092	0,007
1.0-<2.0	0,578	0,024
2.0-<High	0,668	0,026
Total IV	0,057	

Fuente: Elaboración Propia.

**Gráfico N° 13 "WOE CHAR\_220"**



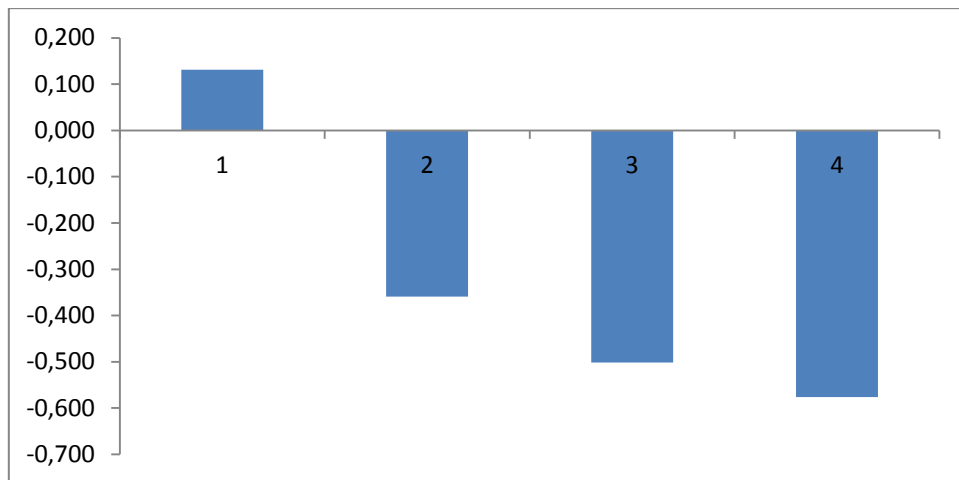
Fuente: Elaboración Propia.

**Tabla N° 30 "WOE CHAR\_226"**

CHAR_226	WOE	Contrib. IV
0.0-<1.0	0,131	0,014
1.0-<22436.0	-0,359	0,007
22436.0-<66812.0	-0,501	0,017
66812.0-<High	-0,576	0,028
<b>Total IV</b>	<b>0,066</b>	

Fuente: Elaboración Propia.

**Gráfico N° 14 "WOE CHAR\_226"**



Fuente: Elaboración Propia.

### Anexos N°3 “Puntajes por variable”

**Tabla N° 31 "Puntajes Modelo de Prueba N°1"**

<b>CHAR_216</b>	<b>Puntaje</b>
0.0-<77938.0	41
77938.0-<242410.0	47
242410.0-<327542.0	51
327542.0-<477921.0	54
477921.0-<915741.0	59
915741.0-<High	66
<b>CHAR_234</b>	<b>Puntaje</b>
0.0-<2.0	45
2.0-<5.0	52
5.0-<19.0	51
19.0-<25.0	49
25.0-<37.0	46
37.0-<High	45
<b>CHAR_222</b>	<b>Puntaje</b>
0.0-<1.0	52
1.0-<2.0	47
2.0-<3.0	45
3.0-<High	43
<b>CHAR_226</b>	<b>Puntaje</b>
0.0-<1.0	51
1.0-<22436.0	48
22436.0-<66812.0	47
66812.0-<High	46
<b>CHAR_220</b>	<b>Puntaje</b>
0.0-<1.0	51
1.0-<2.0	42
2.0-<High	40
<b>CHAR_218</b>	<b>Puntaje</b>
0.0-<1.0	49
1.0-<3.0	54
3.0-<High	55
<b>CHAR_228</b>	<b>Puntaje</b>
0.0-<1.0	49
1.0-<2.0	56
2.0-<High	57
<b>CHAR_230</b>	<b>Puntaje</b>
0.0-<1.0	51
1.0-<489920.0	45
489920.0-<884950.0	49
884950.0-<High	59
<b>CHAR_232</b>	<b>Puntaje</b>
0.0-<1.0	42
1.0-<4.0	53
4.0-<25.0	49
25.0-<High	41
<b>CHAR_224</b>	<b>Puntaje</b>
0.0-<1.0	48
1.0-<2.0	55
2.0-<3.0	58
3.0-<High	60

Fuente: Elaboración Propia.

**Tabla N° 32 "Puntajes Modelo de Pruebas N°2"**

<b>CHAR_216</b>	<b>Puntaje</b>
0.0-<77938.0	43
77938.0-<242410.0	45
242410.0-<327542.0	48
327542.0-<477921.0	51
477921.0-<915741.0	58
915741.0-<High	69
<b>CHAR_234</b>	<b>Puntaje</b>
0.0-<2.0	57
2.0-<5.0	56
5.0-<19.0	49
19.0-<25.0	43
25.0-<37.0	36
37.0-<High	33
<b>CHAR_222</b>	<b>Puntaje</b>
0.0-<1.0	52
1.0-<2.0	46
2.0-<3.0	43
3.0-<High	42
<b>CHAR_226</b>	<b>Puntaje</b>
0.0-<1.0	51
1.0-<22436.0	49
22436.0-<66812.0	49
66812.0-<High	49
<b>CHAR_220</b>	<b>Puntaje</b>
0.0-<1.0	51
1.0-<2.0	42
2.0-<High	40
<b>CHAR_218</b>	<b>Puntaje</b>
0.0-<1.0	49
1.0-<3.0	53
3.0-<High	53
<b>CHAR_228</b>	<b>Puntaje</b>
0.0-<1.0	49
1.0-<2.0	55
2.0-<High	55
<b>CHAR_230</b>	<b>Puntaje</b>
0.0-<1.0	51
1.0-<489920.0	41
489920.0-<884950.0	43
884950.0-<High	50
<b>CHAR_232</b>	<b>Puntaje</b>
0.0-<1.0	42
1.0-<4.0	53
4.0-<25.0	46
25.0-<High	37
<b>CHAR_224</b>	<b>Puntaje</b>
0.0-<1.0	48
1.0-<2.0	53
2.0-<3.0	55
3.0-<High	57

Fuente: Elaboración Propia.

**Tabla N° 33 "Puntajes Modelo Mixto N°1"**

<b>CHAR_200</b>	<b>Puntaje</b>
0.0-<1.0	27
1.0-<2.0	29
2.0-<3.0	42
3.0-<4.0	50
4.0-<5.0	54
5.0-<High	59
<b>CHAR_202</b>	<b>Puntaje</b>
0.0-<8.0	35
8.0-<36.0	46
36.0-<66.0	50
66.0-<7.0E7	58
7.0E7-<High	49
<b>CHAR_204</b>	<b>Puntaje</b>
0.0-<2.0	4
2.0-<5.0	25
5.0-<9.0	32
9.0-<13.0	39
13.0-<High	55
<b>CHAR_208</b>	<b>Puntaje</b>
0.0-<20.0	33
20.0-<49.0	41
49.0-<75.0	47
75.0-<130.0	51
130.0-<1000.0	55
1000.0-<High	52
<b>CHAR_210</b>	<b>Puntaje</b>
0.0-<1.0	44
1.0-<12.0	51
12.0-<44.0	48
44.0-<68.0	44
68.0-<86.0	42
86.0-<1000000.0	39
1000000.0-<High	27

<b>CHAR_212</b>	<b>Puntaje</b>
0.0-<1.0	46
1.0-<44.0	50
44.0-<91.0	49
91.0-<7.0E7	44
7.0E7-<High	39
<b>CHAR_214</b>	<b>Puntaje</b>
0.0-<1.0	47
1.0-<High	31
<b>CHAR_216</b>	<b>Puntaje</b>
1.0-<77938.0	36
77938.0-<242410.0	43
242410.0-<327542.0	47
327542.0-<477921.0	50
477921.0-<915741.0	55
915741.0-<High	62
<b>CHAR_234</b>	<b>Puntaje</b>
0.0-<2.0	51
2.0-<5.0	48
5.0-<19.0	47
19.0-<25.0	45
25.0-<37.0	41
37.0-<High	40
<b>CHAR_222</b>	<b>Puntaje</b>
0.0-<1.0	48
1.0-<2.0	41
2.0-<3.0	39
3.0-<High	37
<b>CHAR_224</b>	<b>Puntaje</b>
0.0-<1.0	44
1.0-<2.0	49
2.0-<3.0	52
3.0-<High	53

Fuente: Elaboración Propia.



**Tabla N° 34 "Puntajes Modelo Mixto N°2"**

<b>CHAR_200</b>	<b>Puntaje</b>
0.0-<1.0	42
1.0-<2.0	33
2.0-<3.0	42
3.0-<4.0	48
4.0-<5.0	51
5.0-<High	56
<b>CHAR_202</b>	<b>Puntaje</b>
0.0-<8.0	40
8.0-<36.0	47
36.0-<66.0	53
66.0-<7.0E7	51
7.0E7-<High	44
<b>CHAR_204</b>	<b>Puntaje</b>
0.0-<2.0	40
2.0-<5.0	47
5.0-<9.0	42
9.0-<13.0	43
13.0-<High	44
<b>CHAR_208</b>	<b>Puntaje</b>
0.0-<20.0	35
20.0-<49.0	42
49.0-<75.0	46
75.0-<130.0	50
130.0-<1000.0	53
1000.0-<High	48
<b>CHAR_210</b>	<b>Puntaje</b>
0.0-<1.0	46
1.0-<12.0	52
12.0-<44.0	46
44.0-<68.0	42
68.0-<86.0	39
86.0-<1000000.0	36
1000000.0-<High	26

<b>CHAR_212</b>	<b>Puntaje</b>
0.0-<1.0	47
1.0-<44.0	47
44.0-<91.0	46
91.0-<7.0E7	44
7.0E7-<High	42
<b>CHAR_214</b>	<b>Puntaje</b>
0.0-<1.0	48
1.0-<High	4
<b>CHAR_216</b>	<b>Puntaje</b>
1.0-<77938.0	39
77938.0-<242410.0	40
242410.0-<327542.0	43
327542.0-<477921.0	46
477921.0-<915741.0	53
915741.0-<High	63
<b>CHAR_234</b>	<b>Puntaje</b>
0.0-<2.0	52
2.0-<5.0	51
5.0-<19.0	44
19.0-<25.0	39
25.0-<37.0	32
37.0-<High	28
<b>CHAR_222</b>	<b>Puntaje</b>
0.0-<1.0	47
1.0-<2.0	41
2.0-<3.0	39
3.0-<High	37
<b>CHAR_224</b>	<b>Puntaje</b>
0.0-<1.0	44
1.0-<2.0	48
2.0-<3.0	50
3.0-<High	51

Fuente: Elaboración Propia.

Anexos N°4 “Listado Puntaje Final Obtenido”

Tabla N° 35 "Puntaje Final Modelo de Pruebas N°1"

Tabla de contingencia SCORE_BHV1 * BUENO_MALO				
Recuento				
		BUENO_MALO		Total
		,00	1,00	
SCORE_BHV 1	410,00	1	0	1
	413,00	5	3	8
	414,00	3	0	3
	415,00	2	0	2
	416,00	1	3	4
	417,00	4	3	7
	418,00	3	4	7
	419,00	6	1	7
	420,00	6	2	8
	421,00	8	1	9
	422,00	3	1	4
	423,00	11	0	11
	424,00	37	0	37
	425,00	13	1	14
	426,00	13	2	15
	427,00	32	11	43
	428,00	15	3	18
	429,00	45	7	52
	430,00	39	3	42
	431,00	27	2	29
	432,00	64	17	81
	433,00	27	3	30
	434,00	29	4	33
	435,00	43	9	52
	436,00	43	5	48
	437,00	37	4	41
	438,00	411	37	448
	439,00	27	5	32
	440,00	18	4	22
	441,00	124	24	148
	442,00	55	3	58
	443,00	107	17	124
	444,00	130	23	153
	445,00	81	15	96
	446,00	123	15	138
	447,00	35	9	44
	448,00	37	4	41
	449,00	52	7	59
	450,00	106	13	119
	451,00	58	6	64
	452,00	84	11	95
	453,00	87	11	98
	454,00	29	8	37
	455,00	50	7	57
	456,00	85	13	98
457,00	138	27	165	
458,00	225	42	267	
459,00	288	61	349	
460,00	171	41	212	
461,00	225	53	278	
462,00	434	81	515	
463,00	802	195	997	
464,00	1619	299	1918	
465,00	1153	235	1388	
466,00	773	134	907	
467,00	1382	207	1589	
468,00	1714	280	1994	
469,00	2328	382	2710	
470,00	3941	626	4567	
471,00	3100	513	3613	
472,00	3386	475	3861	
473,00	3975	585	4560	
474,00	7007	1041	8048	
475,00	10083	1422	11505	

	476.00	7804	1110	8914
	477.00	6744	915	7659
	478.00	11090	1363	12453
	479.00	10174	1373	11547
	480.00	15366	1632	16998
	481.00	20200	2361	22561
	482.00	17262	1994	19256
	483.00	173671	23155	196826
	484.00	26647	2602	29249
	485.00	14751	1419	16170
	486.00	24544	2596	27140
	487.00	50488	3908	54396
	488.00	52327	4636	56963
	489.00	110614	7733	118347
	490.00	44675	3510	48185
	491.00	19018	1399	20417
	492.00	60662	3861	64523
	493.00	46658	3709	50367
	494.00	79038	4952	83990
	495.00	133422	9299	142721
	496.00	92395	5568	97963
	497.00	62457	3053	65510
	498.00	42497	2177	44674
	499.00	49694	2364	52058
	500.00	56802	3048	59850
	501.00	104251	4893	109144
	502.00	78782	3322	82104
	503.00	59928	2334	62262
	504.00	71163	2679	73842
	505.00	43339	1557	44896
	506.00	43460	1545	45005
	507.00	74779	2755	77534
	508.00	69089	2503	71592
	509.00	44956	1215	46171
	510.00	43064	1094	44158
	511.00	31608	852	32460
	512.00	28048	725	28773
	513.00	37087	1025	38112
	514.00	35303	929	36232
	515.00	37405	730	38135
	516.00	29515	539	30054
	517.00	15194	302	15496
	518.00	15309	294	15603
	519.00	16592	292	16884
	520.00	16172	328	16500
	521.00	12136	225	12361
	522.00	10033	158	10191
	523.00	9947	119	10066
	524.00	8511	101	8612
	525.00	4686	77	4763
	526.00	4496	81	4577
	527.00	4910	72	4982
	528.00	4073	42	4115
	529.00	1330	16	1346
	530.00	1124	22	1146
	531.00	916	5	921
	532.00	488	8	496
	533.00	481	10	491
	534.00	332	7	339
	535.00	185	2	187
	536.00	73	4	77
	537.00	27	0	27
	538.00	33	0	33
	539.00	14	0	14
	540.00	11	0	11
	541.00	6	0	6
	542.00	3	0	3
Total		2232819	137619	2370438

Fuente: Elaboración Propia.

Tabla N° 36 "Puntaje Final Modelo de Pruebas N°2"

Tabla de contingencia SCORE_BHV1 * BUENO_MALO				
Recuento				
		BUENO_MALO		Total
		,00	1,00	
SCORE_BHV 1	406,00	1	0	1
	408,00	3	1	4
	409,00	0	2	2
	410,00	0	4	4
	411,00	2	3	5
	412,00	1	1	2
	413,00	4	3	7
	414,00	5	3	8
	415,00	2	3	5
	416,00	4	3	7
	417,00	2	0	2
	418,00	3	10	13
	419,00	7	5	12
	420,00	14	15	29
	421,00	8	10	18
	422,00	10	9	19
	423,00	33	17	50
	424,00	17	21	38
	425,00	23	13	36
	426,00	20	10	30
	427,00	30	14	44
	428,00	15	10	25
	429,00	17	25	42
	430,00	22	10	32
	431,00	38	20	58
	432,00	39	14	53
	433,00	71	57	128
	434,00	18	12	30
	435,00	28	13	41
	436,00	102	50	152
	437,00	57	45	102
	438,00	36	30	66
	439,00	116	69	185
	440,00	69	27	96
	441,00	62	53	115
	442,00	50	49	99
	443,00	104	80	184
	444,00	168	128	296
	445,00	152	79	231
	446,00	634	342	976
	447,00	198	171	369
	448,00	270	163	433
449,00	296	214	510	
450,00	448	359	807	
451,00	387	301	688	
452,00	439	300	739	
453,00	472	311	783	
454,00	958	722	1680	
455,00	1388	778	2166	
456,00	2062	1098	3160	
457,00	2548	1385	3933	
458,00	1475	741	2216	
459,00	2590	1303	3893	
460,00	2180	1161	3341	
461,00	3075	1561	4636	
462,00	3368	1611	4979	
463,00	4100	1928	6028	
464,00	3689	1713	5402	
465,00	6729	2748	9477	
466,00	6085	2558	8643	
467,00	10489	4594	15083	
468,00	6641	2594	9235	
469,00	9832	4155	13987	
470,00	12052	4201	16253	
471,00	10161	3612	13773	
472,00	23952	7995	31947	

473.00	16802	5702	22504	
474.00	15415	4848	20263	
475.00	16578	5083	21661	
476.00	20774	6941	27715	
477.00	19130	4988	24118	
478.00	17775	4673	22448	
479.00	29239	7573	36812	
480.00	54429	11959	66388	
481.00	19000	4290	23290	
482.00	122266	25606	147872	
483.00	17098	3830	20928	
484.00	23040	4310	27350	
485.00	50340	9124	59464	
486.00	50153	8827	58980	
487.00	89351	14811	104162	
488.00	60431	9640	70071	
489.00	115659	18137	133796	
490.00	25177	3975	29152	
491.00	57396	8138	65534	
492.00	57448	8321	65769	
493.00	52024	7104	59128	
494.00	48085	6082	54167	
495.00	95294	11138	106432	
496.00	64365	7577	71942	
497.00	38528	4275	42803	
498.00	97515	9527	107042	
499.00	55226	5744	60970	
500.00	38422	3608	42030	
501.00	57177	5164	62341	
502.00	61404	5205	66609	
503.00	44567	4280	48847	
504.00	90269	7957	98226	
505.00	75111	6278	81389	
506.00	37650	2941	40591	
507.00	34898	2460	37358	
508.00	67849	4716	72565	
509.00	33436	2320	35756	
510.00	36822	2422	39244	
511.00	59777	3675	63452	
512.00	25121	1611	26732	
513.00	27033	1424	28457	
514.00	41393	2232	43625	
515.00	39750	1831	41581	
516.00	15745	780	16525	
517.00	29165	1339	30504	
518.00	22726	1154	23880	
519.00	11516	549	12065	
520.00	23119	872	23991	
521.00	24761	1015	25776	
522.00	14654	568	15222	
523.00	12181	444	12625	
524.00	22252	703	22955	
525.00	3267	143	3410	
526.00	8072	259	8331	
527.00	7600	296	7896	
528.00	10011	363	10374	
529.00	3248	124	3372	
530.00	13828	337	14165	
531.00	6399	178	6577	
532.00	893	14	907	
533.00	2993	99	3092	
534.00	404	4	408	
535.00	1217	46	1263	
536.00	198	5	203	
537.00	5379	129	5508	
538.00	635	9	644	
539.00	14	0	14	
540.00	6	0	6	
541.00	11	0	11	
542.00	4	0	4	
544.00	1	0	1	
548.00	5	1	6	
549.00	2	0	2	
Total		2457394	339315	2796709

Fuente: Elaboración Propia.

**Tabla N° 37 "puntaje Final Modelo Mixto N°1"**

<b>Tabla de contingencia SCORE_BHV * BUENO_MALO</b>				
Recuento				
		BUENO_MALO		Total
		,00	1,00	
SCORE_BHV	282,00	0	1	1
	284,00	2	10	12
	286,00	1	5	6
	288,00	0	12	12
	289,00	1	2	3
	291,00	1	11	12
	292,00	0	1	1
	293,00	3	16	19
	294,00	3	14	17
	295,00	10	56	66
	296,00	5	41	46
	297,00	3	8	11
	298,00	7	25	32
	299,00	3	24	27
	300,00	3	26	29
	301,00	8	59	67
	302,00	11	57	68
	303,00	35	262	297
	304,00	6	16	22
	305,00	11	103	114
	306,00	12	63	75
	307,00	15	141	156
	308,00	27	111	138
	309,00	14	107	121
	310,00	36	209	245
	311,00	12	63	75
	312,00	41	254	295
	313,00	36	141	177
	314,00	111	458	569
	315,00	21	154	175
	316,00	76	335	411
	317,00	77	240	317
	318,00	73	300	373
	319,00	83	356	439
	320,00	62	253	315
	321,00	77	342	419
	322,00	62	229	291
	323,00	122	435	557
	324,00	119	312	431
	325,00	144	507	651
	326,00	157	458	615
	327,00	325	806	1131
	328,00	152	479	631
	329,00	144	343	487
	330,00	221	592	813
	331,00	207	561	768
	332,00	224	517	741
	333,00	277	622	899
	334,00	425	809	1234
	335,00	318	895	1213
	336,00	211	494	705
	337,00	340	721	1061
	338,00	449	722	1171
	339,00	412	1030	1442
	340,00	432	770	1202
	341,00	524	868	1392
	342,00	563	859	1422
	343,00	390	776	1166
	344,00	568	801	1369
	345,00	508	767	1275
	346,00	597	958	1555
	347,00	687	1043	1730
	348,00	899	1107	2006
	349,00	669	856	1525
	350,00	719	947	1666
	351,00	799	1046	1845

	352,00	940	1080	2020
	353,00	634	762	1396
	354,00	975	1135	2110
	355,00	1342	1351	2693
	356,00	934	906	1840
	357,00	963	1033	1996
	358,00	1213	1116	2329
	359,00	1385	1196	2581
	360,00	940	871	1811
	361,00	1316	1028	2344
	362,00	1629	1272	2901
	363,00	1770	1246	3016
	364,00	1185	899	2084
	365,00	1654	1105	2759
	366,00	1695	1109	2804
	367,00	1720	1048	2768
	368,00	1854	1105	2959
	369,00	2252	1199	3451
	370,00	2218	1146	3364
	371,00	2154	1177	3331
	372,00	2563	1151	3714
	373,00	2243	1043	3286
	374,00	2602	1159	3761
	375,00	2663	1033	3696
	376,00	3643	1392	5035
	377,00	2866	1134	4000
	378,00	3656	1330	4986
	379,00	3299	1071	4370
	380,00	3998	1218	5216
	381,00	3595	1146	4741
	382,00	4620	1255	5875
	383,00	4810	1289	6099
	384,00	5042	1231	6273
	385,00	5621	1334	6955
	386,00	5442	1329	6771
	387,00	5551	1209	6760
	388,00	5872	1173	7045
	389,00	6993	1395	8388
	390,00	7994	1664	9658
	391,00	7569	1341	8910
	392,00	8121	1418	9539
	393,00	8961	1350	10311
	394,00	8504	1299	9803
	395,00	9515	1307	10822
	396,00	10050	1383	11433
	397,00	11511	1476	12987
	398,00	11770	1406	13176
	399,00	13027	1425	14452
	400,00	13348	1678	15026
	401,00	14624	1423	16047
	402,00	14895	1502	16397
	403,00	15893	1453	17346
	404,00	15758	1463	17221
	405,00	17663	1588	19251
	406,00	22128	1716	23844
	407,00	18604	1390	19994
	408,00	22683	1607	24290
	409,00	21352	1528	22880
	410,00	22258	1503	23761
	411,00	22604	1230	23834
	412,00	22017	1321	23338
	413,00	30313	1663	31976
	414,00	31676	1618	33294
	415,00	24849	1142	25991
	416,00	26281	1372	27653
	417,00	33367	1527	34894
	418,00	32118	1229	33347
	419,00	31331	1279	32610
	420,00	36566	1352	37918
	421,00	34559	1250	35809
	422,00	41323	1292	42615
	423,00	34103	1201	35304
	424,00	35616	1122	36738
	425,00	41233	1187	42420
	426,00	35712	1011	36723
	427,00	39797	1079	40876
	428,00	43936	1009	44945
	429,00	44181	1063	45244

	430,00	38024	851	38875
	431,00	45321	1008	46329
	432,00	38583	858	39441
	433,00	48307	930	49237
	434,00	41740	758	42498
	435,00	45889	738	46627
	436,00	43536	700	44236
	437,00	41670	659	42329
	438,00	43928	671	44599
	439,00	45517	591	46108
	440,00	41342	530	41872
	441,00	31304	386	31690
	442,00	45175	518	45693
	443,00	50632	556	51188
	444,00	34636	365	35001
	445,00	28858	308	29166
	446,00	37428	308	37736
	447,00	50705	451	51156
	448,00	32859	265	33124
	449,00	20168	165	20333
	450,00	33472	254	33726
	451,00	42083	273	42356
	452,00	33622	225	33847
	453,00	11927	78	12005
	454,00	28367	147	28514
	455,00	32380	153	32533
	456,00	27451	134	27585
	457,00	11174	54	11228
	458,00	15770	75	15845
	459,00	28974	134	29108
	460,00	15296	70	15366
	461,00	7401	30	7431
	462,00	9918	32	9950
	463,00	19864	58	19922
	464,00	10832	43	10875
	465,00	4484	11	4495
	466,00	5082	8	5090
	467,00	9436	21	9457
	468,00	6765	12	6777
	469,00	1641	9	1650
	470,00	1494	2	1496
	471,00	6866	13	6879
	472,00	3091	11	3102
	473,00	790	0	790
	474,00	454	0	454
	475,00	1887	1	1888
	476,00	1155	3	1158
	477,00	129	1	130
	478,00	19	0	19
	479,00	241	1	242
	480,00	898	1	899
	481,00	2	0	2
	482,00	2	0	2
	483,00	6	0	6
	484,00	12	0	12
	<b>Total</b>	<b>2232819</b>	<b>137619</b>	<b>2370438</b>

**Fuente: Elaboración Propia.**



Tabla N° 38 "Puntaje Final Modelo Mixto N°2"

Tabla de contingencia SCORE_BHV * BUENO_MALO				
Recuento				
		BUENO_MALO		Total
		,00	1,00	
SCORE_BH V	362,00	0	1	1
	363,00	0	2	2
	365,00	0	2	2
	366,00	0	1	1
	367,00	0	1	1
	368,00	0	1	1
	369,00	0	1	1
	370,00	0	2	2
	371,00	0	1	1
	372,00	2	4	6
	373,00	3	13	16
	374,00	2	6	8
	375,00	1	15	16
	376,00	5	16	21
	377,00	1	5	6
	378,00	0	12	12
	379,00	0	16	16
	380,00	0	32	32
	381,00	4	21	25
	382,00	2	38	40
	383,00	5	75	80
	384,00	2	45	47
	385,00	9	66	75
	386,00	3	54	57
	387,00	6	66	72
	388,00	4	63	67
	389,00	6	112	118
	390,00	9	106	115
	391,00	14	103	117
	392,00	14	158	172
	393,00	10	122	132
	394,00	16	151	167
	395,00	27	246	273
	396,00	18	158	176
	397,00	19	139	158
	398,00	42	265	307
	399,00	35	228	263
	400,00	15	189	204
	401,00	47	315	362
	402,00	46	305	351
	403,00	26	236	262
	404,00	48	317	365
405,00	76	453	529	
406,00	41	247	288	
407,00	42	313	355	
408,00	54	425	479	
409,00	33	323	356	
410,00	59	386	445	
411,00	97	513	610	
412,00	98	474	572	
413,00	76	350	426	
414,00	76	494	570	
415,00	78	406	484	
416,00	71	463	534	
417,00	125	563	688	
418,00	127	517	644	
419,00	114	484	598	
420,00	165	674	839	
421,00	163	556	719	
422,00	116	455	571	
423,00	162	647	809	
424,00	160	641	801	
425,00	129	528	657	
426,00	197	765	962	
427,00	263	857	1120	
428,00	205	630	835	

	429,00	214	649	863
	430,00	231	754	985
	431,00	230	737	967
	432,00	236	719	955
	433,00	305	961	1266
	434,00	383	1135	1518
	435,00	310	900	1210
	436,00	318	956	1274
	437,00	404	985	1389
	438,00	412	977	1389
	439,00	356	993	1349
	440,00	502	1128	1630
	441,00	599	1272	1871
	442,00	572	1236	1808
	443,00	652	1318	1970
	444,00	648	1244	1892
	445,00	693	1279	1972
	446,00	771	1358	2129
	447,00	867	1472	2339
	448,00	913	1488	2401
	449,00	1083	1674	2757
	450,00	1221	1685	2906
	451,00	1130	1467	2597
	452,00	1217	1584	2801
	453,00	1741	1977	3718
	454,00	1302	1586	2888
	455,00	1415	1667	3082
	456,00	2100	2106	4206
	457,00	1822	1744	3566
	458,00	1607	1666	3273
	459,00	2599	2165	4764
	460,00	2756	2222	4978
	461,00	1936	1740	3676
	462,00	2666	2150	4816
	463,00	2896	2238	5134
	464,00	2642	1938	4580
	465,00	3563	2441	6004
	466,00	3614	2455	6069
	467,00	3432	2245	5677
	468,00	4116	2406	6522
	469,00	4431	2549	6980
	470,00	3819	2174	5993
	471,00	5548	2745	8293
	472,00	5399	2785	8184
	473,00	5125	2615	7740
	474,00	6363	2942	9305
	475,00	7721	3339	11060
	476,00	5932	2694	8626
	477,00	6561	2920	9481
	478,00	8498	3482	11980
	479,00	8632	3284	11916
	480,00	8930	3528	12458
	481,00	10882	3988	14870
	482,00	12398	4046	16444
	483,00	10546	3566	14112
	484,00	12675	4261	16936
	485,00	12828	3943	16771
	486,00	16261	4207	20468
	487,00	15278	4510	19788
	488,00	17199	4713	21912
	489,00	20530	5095	25625
	490,00	20200	5182	25382
	491,00	20465	5408	25873
	492,00	21772	4740	26512
	493,00	27447	5726	33173
	494,00	26657	5417	32074
	495,00	28448	5564	34012
	496,00	31943	6030	37973
	497,00	36121	6343	42464
	498,00	32577	5886	38463
	499,00	36108	6005	42113
	500,00	41159	6166	47325
	501,00	44614	6290	50904
	502,00	43678	5921	49599
	503,00	49522	6301	55823
	504,00	52672	6458	59130
	505,00	52240	5974	58214
	506,00	49647	5685	55332

507,00	53542	5975	59517
508,00	60131	6078	66209
509,00	55702	5247	60949
510,00	60753	5497	66250
511,00	63397	5595	68992
512,00	65281	5132	70413
513,00	57749	4291	62040
514,00	66708	4962	71670
515,00	63054	4396	67450
516,00	60277	3973	64250
517,00	61457	3870	65327
518,00	61705	3702	65407
519,00	63111	3290	66401
520,00	52928	2853	55781
521,00	63120	3210	66330
522,00	59751	2678	62429
523,00	47913	2282	50195
524,00	47967	2206	50173
525,00	52092	2069	54161
526,00	43708	1614	45322
527,00	40092	1564	41656
528,00	45053	1633	46686
529,00	37578	1176	38754
530,00	30292	1008	31300
531,00	29958	928	30886
532,00	33565	872	34437
533,00	24057	592	24649
534,00	25032	687	25719
535,00	24213	598	24811
536,00	18448	405	18853
537,00	13879	327	14206
538,00	17166	335	17501
539,00	15990	229	16219
540,00	10748	220	10968
541,00	12120	208	12328
542,00	9425	127	9552
543,00	5084	88	5172
544,00	4377	64	4441
545,00	7544	81	7625
546,00	2620	43	2663
547,00	3893	46	3939
548,00	3570	50	3620
549,00	779	12	791
550,00	1346	10	1356
551,00	114	3	117
552,00	563	9	572
553,00	561	4	565
554,00	82	0	82
555,00	423	4	427
556,00	3	0	3
557,00	2	0	2
558,00	3	0	3
<b>Total</b>	<b>2457394</b>	<b>339315</b>	<b>2796709</b>

**Fuente: Elaboración Propia.**

## Anexos N°5 “Matriz de Correlaciones”

### Tabla N° 39 "Matriz de Correlaciones Modelo de pruebas N°1"

Matriz de correlaciones

		Constant	WOE_CHAR_216	WOE_CHAR_218	WOE_CHAR_220	WOE_CHAR_222	WOE_CHAR_224	WOE_CHAR_226	WOE_CHAR_228	WOE_CHAR_230	WOE_CHAR_232	WOE_CHAR_234
Paso 1	Constant	1,000	,100	,030	,031	,052	,007	,004	,035	,036	,034	,098
	WOE_CHAR_216	,100	1,000	-,112	-,188	-,197	-,019	,015	,051	-,054	-,056	,311
	WOE_CHAR_218	,030	-,112	1,000	-,240	-,120	,009	-,001	,004	,006	-,027	-,037
	WOE_CHAR_220	,031	-,188	-,240	1,000	,004	-,002	,000	-,025	-,007	,004	,005
	WOE_CHAR_222	,052	-,197	-,120	,004	1,000	,004	-,002	-,007	-,001	-,039	-,069
	WOE_CHAR_224	,007	-,019	,009	-,002	,004	1,000	-,986	-,106	,004	,022	-,012
	WOE_CHAR_226	,004	,015	-,001	,000	-,002	-,986	1,000	,070	-,014	-,011	-,001
	WOE_CHAR_228	,035	,051	,004	-,025	-,007	-,106	,070	1,000	-,009	,033	-,033
	WOE_CHAR_230	,036	-,054	,006	-,007	-,001	,004	-,014	-,009	1,000	-,007	-,090
	WOE_CHAR_232	,034	-,056	-,027	,004	-,039	,022	-,011	,033	-,007	1,000	-,530
	WOE_CHAR_234	,098	,311	-,037	,005	-,069	-,012	-,001	-,033	-,090	-,530	1,000

Fuente: Elaboración Propia.

**Tabla N° 40 "Matriz de Correlaciones Modelo de Prueba N°2"**

**Matriz de correlaciones**

		Constant	WOE_CHAR_216	WOE_CHAR_218	WOE_CHAR_220	WOE_CHAR_222	WOE_CHAR_224	WOE_CHAR_226	WOE_CHAR_228	WOE_CHAR_230	WOE_CHAR_232	WOE_CHAR_234
Paso 1	Constant	1,000	,104	,047	,037	,083	,011	,008	,045	,225	,045	,026
	WOE_CHAR_216	,104	1,000	-,140	-,201	-,213	-,020	,026	,087	,015	-,159	,015
	WOE_CHAR_218	,047	-,140	1,000	-,237	-,107	,011	-,001	,001	,001	-,049	-,044
	WOE_CHAR_220	,037	-,201	-,237	1,000	,003	-,002	,000	-,028	-,002	-,002	,010
	WOE_CHAR_222	,083	-,213	-,107	,003	1,000	,004	-,004	-,013	-,022	-,028	-,054
	WOE_CHAR_224	,011	-,020	,011	-,002	,004	1,000	-,975	-,118	,004	,017	-,012
	WOE_CHAR_226	,008	,026	-,001	,000	-,004	-,975	1,000	,064	-,017	,002	,016
	WOE_CHAR_228	,045	,087	,001	-,028	-,013	-,118	,064	1,000	-,015	,026	-,003
	WOE_CHAR_230	,225	,015	,001	-,002	-,022	,004	-,017	-,015	1,000	,061	-,006
	WOE_CHAR_232	,045	-,159	-,049	-,002	-,028	,017	,002	,026	,061	1,000	-,760
	WOE_CHAR_234	,026	,015	-,044	,010	-,054	-,012	,016	-,003	-,006	-,760	1,000

Fuente: Elaboración Propia.

**Tabla N° 41 "Matriz de correlaciones Modelo Mixto N°1 "**

**Matriz de correlaciones**

		Constant	WOE_CHAR_200	WOE_CHAR_202	WOE_CHAR_204	WOE_CHAR_208	WOE_CHAR_210	WOE_CHAR_212	WOE_CHAR_214	WOE_CHAR_216	WOE_CHAR_222	WOE_CHAR_224	WOE_CHAR_234
Paso 1	Constant	1,000	,084	,121	,109	,129	-,149	,034	-,060	,028	,048	,060	,111
	WOE_CHAR_200	,084	1,000	,108	-,104	-,055	-,087	-,063	,024	-,010	,036	-,047	,186
	WOE_CHAR_202	,121	,108	1,000	,029	,080	-,104	-,012	-,159	-,067	-,035	,007	-,335
	WOE_CHAR_204	,109	-,104	,029	1,000	,067	-,087	-,014	-,636	-,055	-,016	,040	,023
	WOE_CHAR_208	,129	-,055	,080	,067	1,000	-,251	-,324	-,030	-,114	-,012	-,042	-,035
	WOE_CHAR_210	-,149	-,087	-,104	-,087	-,251	1,000	-,142	,058	,168	-,053	-,027	-,285
	WOE_CHAR_212	,034	-,063	-,012	-,014	-,324	-,142	1,000	,005	-,095	-,006	,000	,017
	WOE_CHAR_214	-,060	,024	-,159	-,636	-,030	,058	,005	1,000	,020	,014	-,007	,041
	WOE_CHAR_216	,028	-,010	-,067	-,055	-,114	,168	-,095	,020	1,000	-,227	-,020	,263
	WOE_CHAR_222	,048	,036	-,035	-,016	-,012	-,053	-,006	,014	-,227	1,000	,021	-,050
	WOE_CHAR_224	,060	-,047	,007	,040	-,042	-,027	,000	-,007	-,020	-,020	1,000	-,056
	WOE_CHAR_234	,111	,186	-,335	,023	-,035	-,285	,017	,041	,263	-,050	-,056	1,000

Fuente: Elaboración Propia.

**Tabla N° 42 "Matriz de Correlaciones Modelo Mixto N°2"**

**Matriz de correlaciones**

		Constant	WOE_CHAR_200	WOE_CHAR_202	WOE_CHAR_204	WOE_CHAR_208	WOE_CHAR_210	WOE_CHAR_212	WOE_CHAR_214	WOE_CHAR_216	WOE_CHAR_222	WOE_CHAR_224	WOE_CH AR_234
Paso 1	Constant	1,000	,079	,187	,612	,102	-,054	-,011	-,734	,050	,029	,051	,032
	WOE_CHAR_200	,079	1,000	,129	-,064	-,050	,015	-,070	,026	-,059	,023	-,022	,080
	WOE_CHAR_202	,187	,129	1,000	,052	,037	-,083	,045	-,196	,041	-,053	,001	-,200
	WOE_CHAR_204	,612	-,064	,052	1,000	,056	-,131	-,047	-,728	-,041	-,011	,031	,036
	WOE_CHAR_208	,102	-,050	,037	,056	1,000	-,082	-,362	-,023	,053	-,028	-,092	,085
	WOE_CHAR_210	-,054	,015	-,083	-,131	-,082	1,000	-,303	,117	,093	-,055	-,015	-,288
	WOE_CHAR_212	-,011	-,070	,045	-,047	-,362	-,303	1,000	,018	-,123	,005	,023	,004
	WOE_CHAR_214	-,734	,026	-,196	-,728	-,023	,117	,018	1,000	,010	,021	-,004	,014
	WOE_CHAR_216	,050	-,059	,041	-,041	,053	,093	-,123	,010	1,000	-,258	,062	-,150
	WOE_CHAR_222	,029	,023	-,053	-,011	-,028	-,055	,005	,021	-,258	1,000	,012	-,085
	WOE_CHAR_224	,051	-,022	,001	,031	-,092	-,015	,023	-,004	,062	,012	1,000	,110
	WOE_CHAR_234	,032	,080	-,200	,036	,085	-,288	,004	,014	-,150	-,085	,110	1,000

Fuente: Elaboración Propia.

## Anexos N°6 “Medidas de Bondad Conjunta”

**Tabla N° 43 "Medidas De Bondad Modelo de Prueba N°1 "**

Resumen del modelo			
Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	355909,469 <sup>a</sup>	,089	,119

a. La estimación ha finalizado en el número de iteración 4 porque las estimaciones de los parámetros han cambiado en menos de ,001.

**Fuente: Elaboración Propia.**

**Tabla N° 44 "Medidas de Bondad Modelo de Pruebas N°2"**

Resumen del modelo			
Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	882622,141 <sup>a</sup>	,082	,110

a. La estimación ha finalizado en el número de iteración 4 porque las estimaciones de los parámetros han cambiado en menos de ,001.

**Fuente: Elaboración Propia.**



**Tabla N° 45 "Medidas de Bondad Modelo Mixto N°1"**

Resumen del modelo			
Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	260587,276 <sup>a</sup>	,356	,474

a. La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de ,001.

**Fuente: Elaboración Propia.**

**Tabla N° 46 "Medida de Bondad Modelo Mixto N°2"**

Resumen del modelo			
Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	779383,837 <sup>a</sup>	,212	,282

a. La estimación ha finalizado en el número de iteración 5 porque las estimaciones de los parámetros han cambiado en menos de ,001.

**Fuente: Elaboración Propia.**