



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

**AUTOMATIZACIÓN Y MEJORA DE PROCESOS DE LA GESTIÓN TÉCNICA  
DEL SEGURO OBLIGATORIO DE ACCIDENTES PERSONALES PARA  
CONSORCIO SEGUROS GENERALES S.A.**

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

**GONZALO MATIAS ARAYA SEPULVEDA**

PROFESOR GUÍA:  
PATRICIO WOLFF ROJAS

MIEMBROS DE LA COMISIÓN:

ÁNGEL JIMÉNEZ MOLINA  
CRISTIAN JULIO AMDAN  
JORGE ROMERO CASTILLO

SANTIAGO DE CHILE  
2016

## **RESUMEN EJECUTIVO**

Consortio es uno de los mayores conglomerados financieros del país y la mayor aseguradora del mercado. Entre las filiales dedicadas a los seguros, se encuentra Consorcio generales, la filial encargada de los seguros no-vida. Dentro de esta empresa es que se enmarca el proyecto, intentado potenciar el crecimiento para un producto donde se ve un gran potencial tanto en participación de mercado como en rentabilidad.

El objetivo principal del presente proyecto de grado es el aumento del beneficio y marketshare del seguro obligatorio de accidentes personales (SOAP) mediante la implementación de Analítica y la automatización de procesos claves propios de la gestión técnica<sup>1</sup> del producto. Para esto se aplicara la metodología de ingeniería de negocios así como la construcción e implementación de los sistemas TI correspondientes.

La estructura del proyecto fue dividida en las siguientes etapas: realización del análisis estratégico de la empresa, luego se desarrolló la arquitectura de Macro procesos propios de Consorcio Seguro Generales, se diseñaron los procesos claves en BPMN, para finalmente diseñar e implementar los sistemas TI necesarios.

Entre los procesos claves que fueron rediseñados destacan el cálculo de OYNR, en donde se automatiza el cálculo y se mejora considerablemente el poder de predicción del modelo, además de la segmentación y el pricing para la cartera de clientes, así como la creación de un modelo de clasificación de clientes.

Varios resultados relevantes se obtuvieron producto del rediseño y de la implementación de la analítica, por un lado se obtuvo mejoras operacionales en el tiempo de proceso, reducción del orden del 75% y un nuevo modelo de predicción de OYNR con un 93% de efectividad, asimismo se descubrió que variables como la marca del vehículo, su antigüedad o si el seguro fue contratado por una empresa son mejores discriminadores que las variables utilizadas actualmente en la segmentación y pricing.

Finalmente la evaluación económica entrega un VAN de 5300 UF con condiciones conservadoras y un horizonte de 3 años. Luego se discuten cuales son algunos de los mayores desafíos posteriores la implementación y las perspectivas para desarrollos futuros.

---

<sup>1</sup> Gestión Técnica: Gestión propia de las áreas técnicas en las empresas de Seguros. Generalmente relacionada con la suscripción y análisis de riesgos, tarificación, cálculos actuariales, etc.

## **Agradecimientos**

A Maxito, por darme los mejores años de mi vida y a la vez dejarme con un vacío que jamás lograre llenar, no hay día que no piense en ti y rememore tantos momentos magníficos. Nos volveremos a ver.

A mi madre, responsable de cada uno de mis logros y de la persona que me he convertido. Me faltaran días en mi vida para agradecerte por todo el amor que me has dado. Te amo.

A mi familia, las personas más importantes de mi vida, quienes aún al día de hoy no dejan de preocuparse por mí. Gracias a mi abuelita Francisca, a mi tía Ximena y mi tío Rodrigo, nunca podre retribuirles todo lo que han hecho y siguen haciendo por mí.

Al amor de mi vida, mi princesa Guenevere. Sin tu apoyo contante esto no hubiera sido posible, tu amor y preocupación me han dado el aliento y el apoyo cuando lo necesite. Cada día te amo más y me fascina el rumbo en que estamos llevando nuestras vidas.

A mis amigos, a los de la vida y a los de Beauchef. Gracias por todo, cada vez que nos vemos lo pasamos genial y saber que siempre cuento con ustedes me da una gran tranquilidad.

A mis profesores. En especial a mi profesor guía Patricio Wolff, gracias por el tiempo y las valiosas correcciones que le dieron la forma final a esta tesis. Finalmente un saludo a todos los profesores de este maravilloso Magister que me entregaron herramientas invaluable para mi desarrollo académico y profesional.

A través de las puertas del sueño custodiadas por los gules,  
Más allá de los abismos de la noche iluminados por la pálida luna,

He vivido mis vidas sin número.

(Némesis, Lovecraft)

# Tabla de Contenido

Agradecimientos .....	iii
Índice de Figuras.....	viii
Índice de Tablas.....	x
<b>Introducción.....</b>	<b>1</b>
1.1.    Historia de los Seguros .....	1
1.1.1.    Edad antigua .....	1
1.1.2.    Edad Media (hasta mitad del siglo XIV).....	2
1.1.3.    Mediados del siglo XIV a finales del XVII.....	2
1.1.4.    Siglo XVII hasta mediados del XIX .....	2
1.1.5.    Mediados de siglo XIX hasta nuestros días.....	3
1.1.6.    Historia del seguro en Chile .....	3
1.2.    Consortio Seguros.....	4
1.2.1.    Resumen Histórico .....	4
1.2.2.    Filiales .....	6
1.2.3.    SOAP.....	7
1.2.4.    Mercado.....	9
1.3.    Objetivo General.....	13
1.4.    Alcances .....	13
<b>Marco Conceptual.....</b>	<b>14</b>
2.1.    Marco Teórico .....	14
2.1.1.    Bases de datos Relacional .....	14
2.1.1.1.    Data Warehouse.....	16
2.1.2.    Data Mining .....	18
2.1.2.1.    Clustering .....	20
2.1.2.2.    Clasificación .....	21
2.2.    Marco Metodológico.....	24
2.2.1.    Ingeniería de Negocios.....	24
<b>Análisis Estratégico.....</b>	<b>26</b>
3.1.    Misión y Visión .....	26
3.2.    Mapa estratégico.....	26
3.3.    Modelo Delta.....	28
3.4.    Modelo de Negocios .....	29
3.5.    Análisis Mercado .....	30
3.5.1.    FODA.....	30
3.5.2.    Fuerzas de Porter.....	31

<b>Arquitectura de Macro-Procesos .....</b>	<b>33</b>
4.1. Macro-Procesos .....	33
4.2. Variables de diseño .....	38
4.2.1. Estructura Empresa Mercado .....	38
4.2.2. Anticipación.....	38
4.2.3. Coordinación.....	39
4.2.4. Prácticas de Trabajo .....	40
4.2.5. Integración de Procesos Conexos .....	40
4.2.6. Mantención Consolidada de Estado .....	41
4.2.7. Objetivos Específicos .....	41
<b>Diseño de Procesos .....</b>	<b>42</b>
5.1. Diseño BPMN .....	42
5.1.1. Análisis Clientes/Carga DW .....	42
5.1.1.1. Automatización .....	43
5.1.2. Análisis Clientes/Preparación de Modelos .....	47
5.1.2.1. OYNR.....	48
5.1.2.2. Backtesting .....	50
5.1.3. Análisis Clientes/Ejecución de Modelos .....	51
5.1.3.1. Clustering .....	51
5.1.3.2. K-Means .....	52
5.1.3.3. Conglomerado Bietapico .....	54
5.1.4. Venta y Atención Clientes .....	55
5.1.4.1. Clasificación .....	57
5.1.4.2. Redes Neuronales.....	57
5.1.4.3. Árbol de Decisión .....	58
<b>Arquitectura TI.....</b>	<b>62</b>
6.1. Diagramas UML .....	62
6.1.1. Diagrama de Paquetes .....	63
6.1.2. Casos de Uso.....	63
6.1.3. Diagrama de Clases.....	64
6.1.4. Diagrama de secuencia .....	65
6.1.5. Diagrama de despliegue .....	67
<b>Resultados .....</b>	<b>68</b>
7.1. Resultados Segmentación.....	68
7.2. Pricing .....	72
7.3. Prototipo .....	74
7.4. Gestión del Cambio .....	77
7.4.1. Análisis Situación Actual .....	78
7.4.2. Formación de Equipos.....	78
7.4.3. Crear y comunicar visión.....	78

7.4.4.	Avanzar por etapas.....	78
7.4.5.	Desarrollar Habilidades.....	78
7.4.6.	Evaluación y Cierre.....	78
7.4.7.	Roles y estructura organizacional.....	79
7.5.	Evaluación Implementación.....	80
7.6.	Evaluación Económica.....	80
7.6.1.	Estimación Demanda.....	80
7.6.2.	Inversión Inicial.....	82
7.6.3.	Beneficios.....	83
7.6.4.	Costos.....	84
7.6.5.	Flujo de Caja.....	84
	Conclusión.....	89
	Bibliografía.....	93
	Anexos.....	95
	A. Estadísticas completas SOAP-SVS.....	95
	B. Notación BPMN.....	100
	C. Venta Web SOAP.....	105

# Índice de Figuras

Figura 1: Filiales Consorcio.....	6
Figura 2: Patrimonio, Activos y Utilidades Filiales.....	6
Figura 3: Evolución Activos y Patrimonio Filiales.....	7
Figura 4: Ingresos por tipo de producto.....	8
Figura 5: Costo total siniestros.....	9
Figura 6: Número de personas Siniestradas.....	10
Figura 7: Numero de seguros contratados.....	10
Figura 8: Prima Total.....	11
Figura 9: Siniestralidad.....	12
Figura 10: Ejemplo modelo Relacional.....	15
Figura 11: Esquema Data Warehouse.....	18
Figura 12: Esquema Data Mining.....	20
Figura 13: Metodología Ingeniería de Negocios.....	24
Figura 14: Mapa Estratégico CFSA.....	27
Figura 15: Mapa Estratégico Gerencia Técnica.....	28
Figura 16: Modelo de Hax “Mejor Producto”.....	29
Figura 17: Modelo de negocios Consorcio.....	30
Figura 18: Macroprocesos.....	34
Figura 19: Macro 1, Cadena de Valor.....	35
Figura 20: Administración Relación con el Cliente.....	36
Figura 21: Marketing y análisis de mercado.....	37
Figura 22: BP1.....	37
Figura 23: Carga Date Warehouse.....	43
Figura 24: Proceso general en Pentaho.....	44
Figura 25: Validacion de Daton en Pentaho.....	44
Figura 26: Carga DW en Pentaho.....	45
Figura 27: Updates en Pentaho.....	45
Figura 28: Procesamiento de Siniestros eb Pentaho.....	46
Figura 29: Preparacion de Modelos.....	48
Figura 30: Ejemplo Chain Ladder.....	49
Figura 31: Calculo de OYNR (1).....	50
Figura 32: Calculo de OYNR (2).....	50
Figura 33: Ejecucion de Modelos.....	52
Figura 34: Modelo Clustering en RapidMiner.....	53
Figura 35: Comparacion Algoritmos Clustering.....	55
Figura 36: Venta y Atencion Clientes.....	57
Figura 37: Modelo de Clasificacion Red Neuronal RapidMiner.....	58
Figura 38: Modelo de Clasificacion Arbol de Decision RapidMiner.....	60
Figura 39: Diagrama de Paquetes.....	63
Figura 40: Caso de Uso.....	64
Figura 41: Diagrama de Clases.....	65
Figura 42: Diagrama de Secuencia.....	66
Figura 43: Diagrama de Despliegue.....	67

Figura 44: Tipo de Clientes SOAP. ....71

Figura 45: Interfaz Prototipo .....74

Figura 46: Interfaz Prototipo/Validacion Data. ....72

Figura 47: Interfaz Prototipo/OYNR .....75

Figura 48: Interfaz Prototipo/Clustering .....76

Figura 49: Interfaz Prototipo/Clasificacion .....76

Figura 50: Modleo de Gestion de Cambio .....77

Figura 51: Mapa de Roles Proyecto.....79

Figura 52: Demanda de SOAP .....82

## Índice de Tablas

Tabla 1: Estructura Empresa Mercado.....	38
Tabla 2: Anticipación.....	39
Tabla 3: Coordinación.....	39
Tabla 4: Prácticas de Trabajo .....	40
Tabla 5: Integración de Procesos Conexos .....	40
Tabla 6: Mantención Consolidada de Estado .....	41
Tabla 7: Indicadores Backtesting OYNR.....	51
Tabla 8: Indicadores Número de Clusters.....	53
Tabla 9: Distancias entre los centros de los conglomerados finales.....	54
Tabla 10: ANOVA Conglomerados K-Means. ....	54
Tabla 11: Distribucion Categorías Clasificación.....	57
Tabla 12: Resumen Red Neuronal.....	59
Tabla 13: Resumen Árbol de Decisión .....	60
Tabla 14: Sensibilidad y Especificidad Red Neuronal .....	61
Tabla 15: Sensibilidad y Especificidad Árbol de Decisión.....	61
Tabla 16: Características Clusters .....	68
Tabla 17: Distribución geográfica Clusters.....	69
Tabla 18: Antigüedad Promedio vehículos Clusters.....	69
Tabla 19: Siniestralidad Marcas más representativas. ....	70
Tabla 20: Distribución Tipo de Clientes .....	71
Tabla 21: Características Tipos de Clientes .....	71
Tabla 22: Cálculo Prima Inventario.....	73
Tabla 23: Demanda de SOAP.....	80
Tabla 24: Resumen Regresión Demanda (1).....	81
Tabla 25: Resumen Regresión Demanda (2).....	81
Tabla 26: Resumen Regresión Demanda (3).....	81
Tabla 27: Inversión Inicial Proyecto.....	83
Tabla 28: Nuevas Pólizas Por Periodo.....	83
Tabla 29: Ingresos Proyecto.....	83
Tabla 30: Costos Proyecto.....	84
Tabla 31: Flujo de Caja Proyecto (Escenario más probable) .....	86
Tabla 32: Flujo de Caja Proyecto (Escenario Optimista) .....	87
Tabla 33: Flujo de Caja Proyecto (Escenario Pesimista).....	88

# Capítulo 1

## Introducción

### 1.1. Historia de los Seguros

La historia del Seguro se remonta a los albores de la civilización, donde se utilizaban prácticas que constituyeron los inicios de nuestro actual sistema de Seguros. Probablemente las formas más antiguas de Seguros fueron iniciadas por los babilonios y los hindús. Estos primeros contratos eran conocidos bajo el nombre de Contratos a la Gruesa y se efectuaban, esencialmente, entre los banqueros y los propietarios de los barcos. Con frecuencia, el dueño de un barco tomaría prestados los fondos necesarios para comprar carga y financiar un viaje [4].

#### 1.1.1. Edad antigua

En documentos muy antiguos se reflejan vestigios de lo que podrían ser los primeros intentos de la actividad aseguradora, o cuanto menos, de un cierto ambiente de solidaridad promovido por las desgracias que a menudo ocurrían a los componentes de las caravanas que cruzaban la antigua Babilonia, y que a menudo estaban sometidos a robos y pillajes, cuyos daños se soportaban en forma solidaria.

Más adelante, en el siglo V antes de Cristo, existían en Rodas, leyes que contemplaban la obligación que tenían los cargadores al objeto de contribuir a la reparación económica de los daños causados por tempestades, e incluso rescates de buques en poder de los enemigos. En la época de esplendor, y dominación romana se conoce que existieron asociaciones de artesanos, que por medio del pago de una cuota, aseguraban sus propios funerales.

Existían además formas de aseguramiento de buques, generalmente relacionados con la práctica de préstamos a la gruesa, donde el armador a menudo, tomaba a préstamo una cantidad que se correspondía con el valor de la mercadería transportada en su buque. Solo pagaba el capital, (más un elevado tipo de interés), en el supuesto de que arribara al puerto. En caso de naufragio, hacía suyo el citado préstamo. Esta actividad no puede considerarse exactamente un seguro, ya que, el pago de la prima estaba condicionado a la existencia de un siniestro, y la indemnización, se entregaba anticipadamente. [4].

### **1.1.2. Edad Media (hasta mitad del siglo XIV)**

A través de los gremios surgen las llamadas "Guildas", empresas que se parecían a mutualidades de carácter religioso generalmente, y que poco a poco, fueron cobrando aspecto lucrativo, y por lo tanto comercial.

En el siglo XIII, nace el seguro marítimo. En esta época se instaura el sistema de "pago por daño".

Un decreto del Papa Gregorio IX del año 1230, prohíbe el pago de todo tipo de intereses, con lo que se modificaban las operaciones de préstamo a la gruesa ya mencionados. Mediante esta disposición se instaura la prima de seguro, y se abona la indemnización en caso de ocurrir el siniestro. Estas prácticas se acercan más al concepto moderno del seguro. [4].

### **1.1.3. Medios del siglo XIV a finales del XVII**

Aparecen los primeros contratos de seguros marítimos a consecuencia de la creciente actividad de este sector económico.

El primer documento del que se guarda testimonio escrito, pertenece a un contrato de seguro, del buque " Santa Clara", en viaje desde Génova hasta Mallorca, allá por el año 1347. Este contrato ya tomaba la denominación de "Polizza".

Al mismo tiempo aparecen las primeras disposiciones legales de la actividad aseguradora, que de estar representada por una persona individual, pasa a tener la forma de entidades pluripersonales, y sociedades anónimas.

El tremendo incendio que sufrió la ciudad de Londres en 1666, delató la imperiosa necesidad de dar cobertura a las precarias construcciones de la época, por lo tanto es en ese tiempo, cuando se crean las primeras aseguradoras de incendios.

Es en París en 1668 donde se funda lo que se puede considerar la primera Compañía de Seguros por acciones, dedicada al seguro marítimo.

Se crea también, pero en Londres, el famoso Lloyd`s inglés, una corporación de aseguradores particulares. Es conocida su fundación a raíz de las reuniones de aseguradores de riesgos en el café de un tal Edward Lloyd. Esto sucedió en el año 1769. [4].

### **1.1.4. Siglo XVII hasta mediados del XIX**

Es en esta época cuando se da un salto cualitativo, dado que ya no solo se aseguran hechos naturales, sino que aparecen coberturas para daños causados por las personas, y también a las propias personas. Este tipo de seguros obligan a realizar cálculos mucho más técnicos, que los que se hacían hasta ese tiempo. En este momento comienzan a aparecer los conocidos seguros de vida.

Blaise Pascal, y Pierre de Fermat, hicieron trabajos relacionados con el cálculo de probabilidades. Así mismo son conocidos los estudios efectuados por Edmond Halley, sobre la mortalidad, y los realizados por Daniel Bernoulli, en su formulación de "La ley de los grandes números". Sumando todos estos aportes, nace aquí, la ciencia aseguradora.

En el siglo XVIII nacen en Europa, numerosas empresas aseguradoras, similares en su constitución, y estructura a las actuales.

El siglo XIX trae a la actividad, el desarrollo, y la complejidad que carecía. [4].

#### **1.1.5. Medios de siglo XIX hasta nuestros días**

La evolución que experimenta la sociedad durante el siglo XIX, que afecta a las condiciones de vida tiene su influencia en el desarrollo del sector asegurador.

La cada vez más amplia industrialización, y el aumento de la población urbana hacen surgir fenómenos que sin duda favorecen la idea individualista en contraposición con la idea de seguridad del grupo familiar imperante hasta ahora.

Aparecen nuevos riesgos que muchas empresas asumen sin demasiadas garantías, por lo que surge un fuerte intervencionismo por parte de los Estados.

Es en esta época cuando el sector asegurador, como institución, contribuye al fuerte desarrollo industrial, y adapta sus estructuras, a las nuevas exigencias del mercado, y a su vez éste encuentra en el seguro, un gran soporte para su evolución. [4].

#### **1.1.6. Historia del seguro en Chile**

La historia del seguro en Chile comienza a mediados del siglo XIX, época en la cual se hacía patente la necesidad de resguardar, principalmente, el tráfico marítimo. Para satisfacer esta necesidad, británicos radicados en Valparaíso, empezaron a actuar como agentes de compañías de seguros inglesas. [5].

La actividad aseguradora experimentó un crecimiento gradual, apoyada en la regulación del Código del Comercio, dictado en 1865, estableció la normativa básica para el desarrollo de la industria aseguradora. Con el objeto de unificar esfuerzos en pro del desarrollo de esta actividad y contar con una entidad representativa de la opinión del sector, el 4 de septiembre de 1899 se funda la Asociación de Aseguradores de Chile, constituyéndose como la primera en su rubro en Latinoamérica y dando forma a un gremio con una sólida tradición y trayectoria.

Hasta 1980 el mercado asegurador funcionó en un contexto estrictamente regulado, lo que gravitó principalmente en las decisiones de inversión y en las condiciones comerciales del sector.

Con la puesta en vigencia del D.L 3057 en 1980, se aplicaron a este rubro, los principios de subsidiaridad, apertura y competencia, características de la economía social de mercado que provocaron una mayor dinámica en la industria y libertad en el sistema.

Por otra parte, se elevó relevantemente el capital mínimo que permitía operar a una compañía y se establecieron severos límites y resguardos sobre la seguridad, diversificación y liquidez de las inversiones efectuadas por las empresas.

En lo que respecta a las dos últimas décadas, la industria ha experimentado un explosivo desarrollo, impulsado en gran parte por la Reforma Previsional regulada en el D.L.3500, que estableció el régimen privado en pensiones e instituyó las rentas vitalicias previsionales. Hoy, el mercado asegurador chileno es uno de los más importantes inversionistas institucionales del país y cuenta con la participación de importantes grupos aseguradores a nivel mundial.

## **1.2. Consortio Seguros**

CONSORCIO fue formada inicialmente bajo el nombre de Compañía de Seguros La Industrial, en el año 1916. A lo largo de los años y hasta hoy ha sufrido una serie de fusiones con otras compañías menores hasta adquirir en el año 1976 el nombre CONSORCIO Nacional de Seguros (CNS). En 1986 fue comprada por Bankers Trust N.Y.C, una importante institución de la banca mundial. Luego se incorporaron como socios The Bank of Tokio y Chemical Bank N.Y.

Alrededor del año 1997 se incorporaron el grupo Banvida S.A. (Grupo Fernández León y José Antonio Garcés) y P&S (Grupo Hurtado Vicuña), para consolidarse como socios mayoritarios con un 99,76% contribuyendo a que CONSORCIO mantenga su liderazgo como la compañía de seguros más grande del mercado por su volumen total de ventas, activos, patrimonio, inversiones y cobertura geográfica.

CONSORCIO es uno de los mayores conglomerados financieros del país, además de ser el mayor grupo asegurado contando con activos por sobre los 11.400 millones [USD], un patrimonio de 1.370 millones [USD], más de 1.000.000 de clientes y 2.644 colaboradores.

### **1.2.1. Resumen Histórico**

A continuación se resumen algunos de los principales hitos de la compañía [1]:

- 1916: Fundación de la Compañía de Seguros La Industrial.
- 1947: La empresa toma el nombre de Compañía de Seguros de Vida La Industrial.
- 1975: Nuevamente cambia su nombre a Compañía de Seguros de Vida Consortio Nacional de Seguros S.A.

- 1986: La Compañía pasa a manos de Bankers Trust, que se adjudica el 96,21% del total de acciones.
- 1987: Creación de Consorcio Corredores de Bolsa y Consorcio Créditos Hipotecarios.
- 1993: Se inaugura el Edificio Corporativo de Consorcio, diseñado por los destacados arquitectos chilenos Borja Huidobro y Enrique Browne.
- 1999: Consorcio se convierte en una compañía 100% nacional y pasa a manos de las familias Hurtado Vicuña y Fernández León. Nace Consorcio Financiero S.A. (CFSA). CFSA adquiere Cruz Blanca Seguros Generales, la que se transforma en Consorcio Seguros Generales S.A.
- 2000: Se autoriza a operar a Consorcio Créditos Hipotecarios y se crea Consorcio Corredores de Bolsa. Se formaliza asociación con Compass Group N.Y.
- 2002: Se inicia la comercialización de Ahorro Previsional Voluntario (APV). Consorcio se consolida adquiriendo CNA, actual CNLife. Fundación Consorcio inaugura Colegio Monte Olivo en Puente Alto.
- 2002: Inauguración del Colegio Consorcio Vida. 480 niños de escasos recursos de Puente Alto que diariamente reciben educación y alimentación gratuita. Esto permitió que el año 2003, la Fundación Consorcio Vida fuera una de las tres ganadoras del premio "Al Buen Ciudadano empresarial 2003".
- 2003: Consorcio lanza la primera Corredora de Bolsa En Línea. Fundación Consorcio gana premio al Buen Ciudadano Empresarial, entregado por la Cámara Chileno Norteamericana de Comercio.
- 2004: Nace Consorcio Tarjetas de Crédito S.A. y se convierte en el primer emisor no bancario de tarjetas del país.
- 2005: Se crea Consorcio Corredores de Bolsa de Productos. Consorcio es la primera compañía de seguros en ofrecer Créditos de Consumo a pensionados.
- 2006: Consorcio celebra 90 años de historia y es reconocido como Superbrands por el Consejo de Marcas de Chile y Superbrands Internacional.
- 2008: Consorcio lanza el primer portal de venta de seguros de Auto y Viajes en línea.
- 2009: Consorcio compra Banco Monex y crea Banco Consorcio. Compass Consorcio Asset Management firma una alianza entre su filial Consorcio Administradora General de Fondos con LarrainVial. Consorcio obtiene el EFFIE PLATA en su participación en el certamen de los Premios a la Eficiencia Publicitaria Effie Awards 2009, destacando entre un selecto grupo de empresas reconocidas por la efectividad en sus comunicaciones de marketing.

- 2010: Consorcio Financiero reactiva la comercialización de rentas vitalicias a través de su filial CNLife Seguros Vida S.A, una nueva opción para quienes se inclinen por esta modalidad de pensión. Se lanzan las aplicaciones móviles de Consorcio Seguros Vida y Consorcio Corredores de Bolsa, herramienta que introduce mayor tecnología en plataformas a clientes.
- 2011: CNLife obtiene el EFFIE PLATA en su participación en el certamen de los Premios a la Eficiencia Publicitaria Effie Awards 2011.

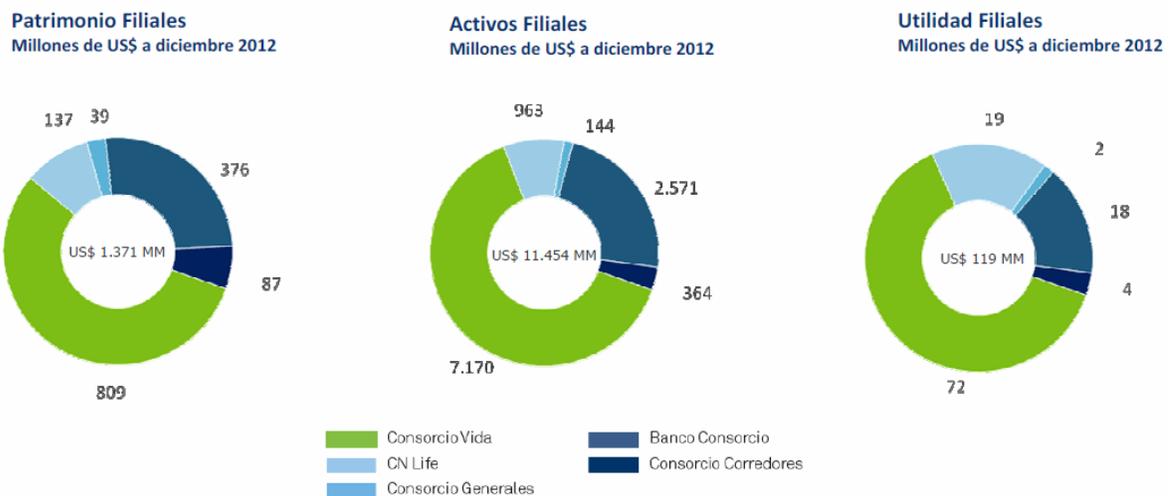
**1.2.2. Filiales**

A través de sus distintas filiales CONSORCIO Financiero es capaz de ofrecer una amplia gama de servicios financieros a la vez que genera una relación permanente y cercana con sus clientes.



**Figura 1: Filiales de Consorcio [2].**

Consorcio vida es la filial más relevante en términos de patrimonio, activos y generación de flujos, aunque vale destacar la participación del banco Consorcio que en pocas años se ha convertido en la segunda filial más importante del conglomerado.



**Figura 2: Patrimonio, Activos y Utilidades Filiales [2].**

En las siguientes graficas se puede apreciar la evolución de los activos y del patrimonio para las distintas filiales de Consortio, nuevamente se reafirma la posición preponderante de Consortio vida y el enorme crecimiento del banco Consortio.

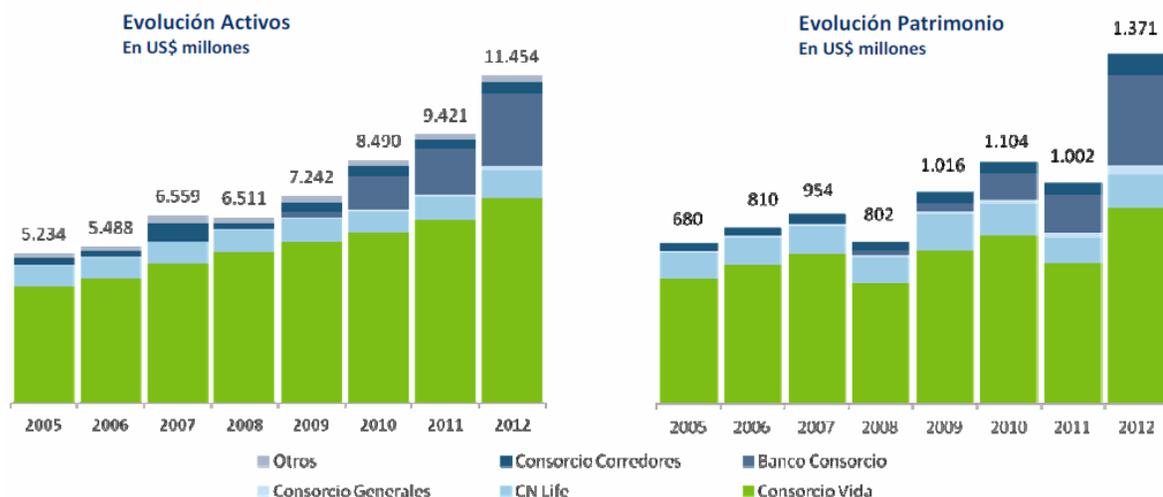


Figura 3: Evolución Activos y Patrimonio Filiales [2].

### 1.2.3. SOAP

A continuación se realiza un breve repaso de en qué consiste el seguro obligatorio de accidentes personales (SOAP) y cuáles son sus características principales.

Seguro Obligatorio de Accidentes Personales (SOAP).

- Es un seguro exigido y regulado por la Ley N° 18.490.
- La póliza del seguro SOAP es única, la cual es elaborada por la SVS.
- Este seguro debe ser contratado por todo propietario de un vehículo motorizado, remolque, acoplados, casas rodantes u otros similares, al momento de adquirir su Permiso de Circulación.
- Se puede contratar directamente en una aseguradora o por internet en los sitios web de éstas.
- No cubre los daños causados al vehículo, sólo ampara los riesgos de muerte y lesiones de las personas a consecuencia de accidentes y reembolsa los gastos médicos y hospitalarios.
- Cubre los gastos médicos relativos a la atención pre-hospitalaria y el transporte sanitario, la hospitalización, la atención médica, quirúrgica, dental, prótesis, implantes, los gastos farmacéuticos y de rehabilitación de las víctimas.

- Su contratación consta en un Certificado que es entregado por la compañía de seguros al propietario del vehículo, el cual contiene entre otras menciones, los datos del vehículo y la individualización de su propietario.
- El seguro SOAP tiene preferencia frente cualquier otra prestación cubierta por una ISAPRE o FONASA u otro seguro complementario de salud o accidente.
- Montos de indemnización:
  - Muerte: 300 UF por persona
  - Incapacidad permanente total: 300 UF por persona
  - Incapacidad permanente parcial: hasta 200 UF por persona
  - Gastos médicos hospitalarios: hasta 300 UF (nivel 03 FONASA)

Este seguro es una de las principales líneas de negocio de Consortio Seguros Generales, además de tener fuertes planes de crecimiento dada la buena rentabilidad de años anteriores y las posibilidades reales de aumentar el porcentaje del Market share que se tiene. En la siguiente figura se puede apreciar la distribución de ingresos por tipo de producto de Generales, SOAP aporta con un 10% de los ingresos de la empresa.

**Ingresos por tipo de producto**  
A diciembre 2012

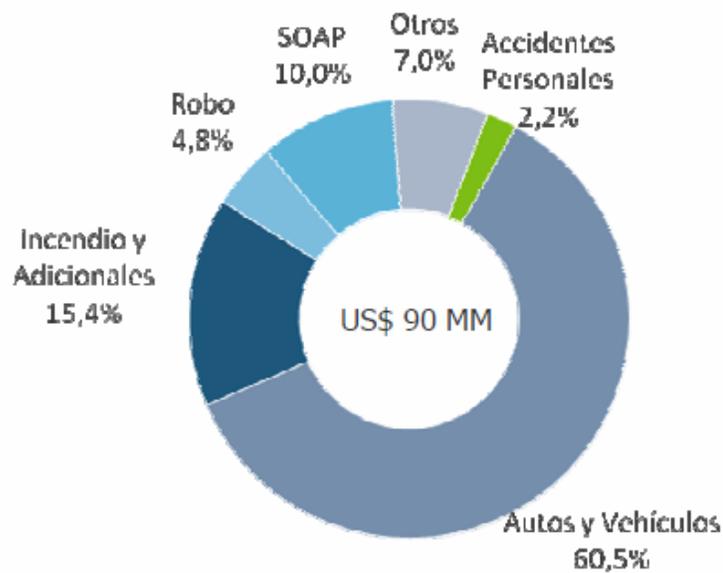


Figura 4: Ingresos por tipo de producto. [2]

**1.2.4. Mercado**

Las siguientes figuras muestran distintas vistas del SOAP para las distintas aseguradoras del mercado chileno. Observando estos datos se puede tener una idea preliminar de la estrategia de Consortio con respecto a este seguro. Los siguientes datos corresponden al periodo de enero a diciembre del 2012 y provienen de la superintendencia de seguro y valores [3]. Las estadísticas completas pueden ser consultadas en el Anexo A.

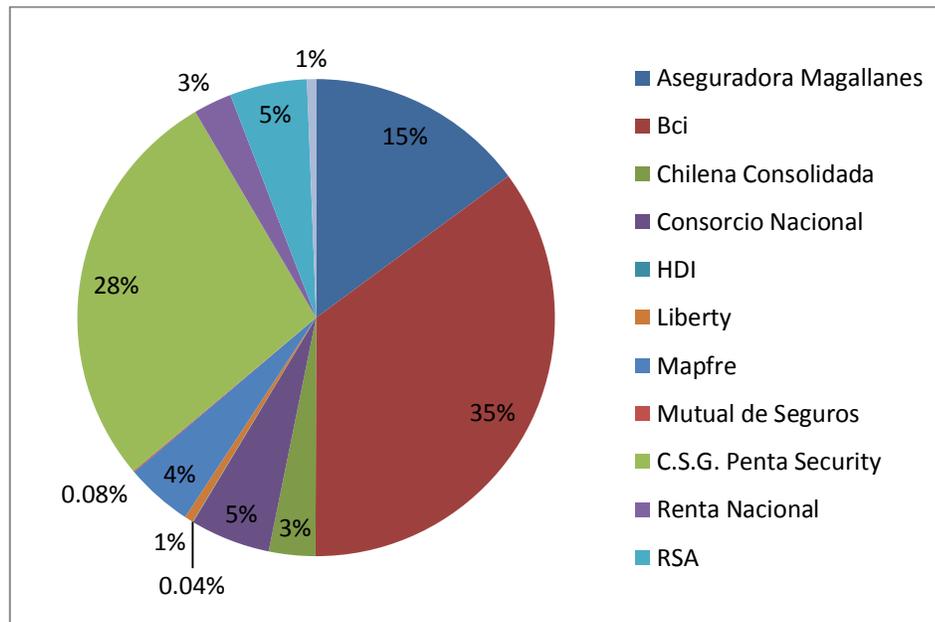
- **Siniestros Pagados Totales:** Corresponden al costo total por siniestros en el periodo en estudio. Se calculan de la siguiente forma

**Costos por Siniestros**

$$= \text{Siniestros pagados} + \text{Siniestros por pagar} + \text{Reserva OYNR}^2 - \text{Costos de siniestros del periodo anterior}$$

Durante el año 2012 el costo total por siniestros para todas las aseguradoras alcanzo el monto de 34.454.702 M\$<sup>3</sup>.

Como se aprecia en la siguiente figura, Consortio se llevó un 5% de los costos totales por siniestros.



**Figura 5: Costo total siniestros. Fuente: SVS**

<sup>2</sup> OYNR = Ocurrido Y No Reportado, Siniestros que han ocurrido a la fecha pero aún no han sido denunciados a la compañía correspondiente.

<sup>3</sup> M\$ = Miles de pesos chilenos.

- Número total de personas siniestradas:** Cada siniestro puede tener uno o más afectados y en el grafico siguiente se puede ver como se distribuyeron el número total de personas siniestradas del año 2012. En total existieron 48.096 personas afectadas.

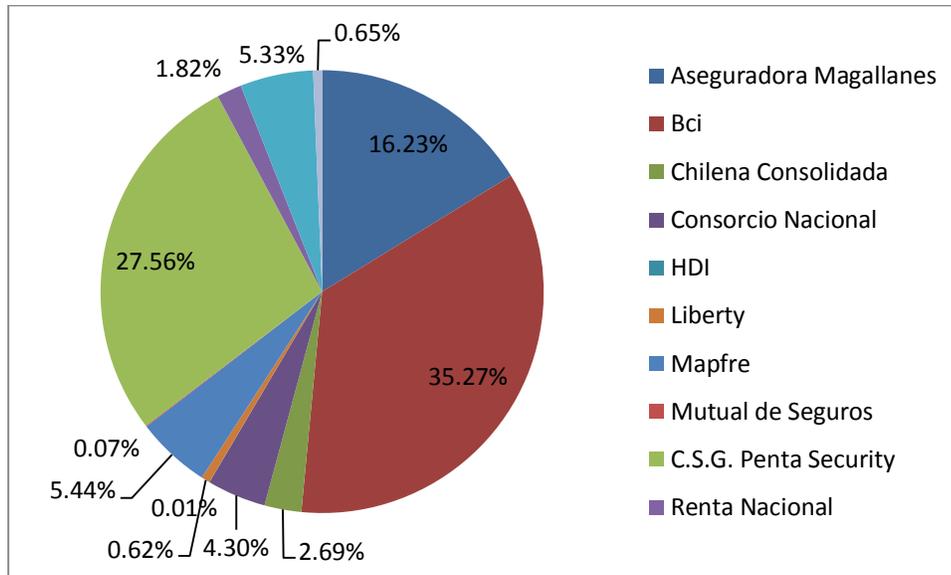


Figura 6: Número de personas Siniestradas. Fuente: SVS

- Numero de seguros totales:** Durante el 2012 se vendieron 4.231.550 seguros y estos fueron distribuidos de la siguiente forma:

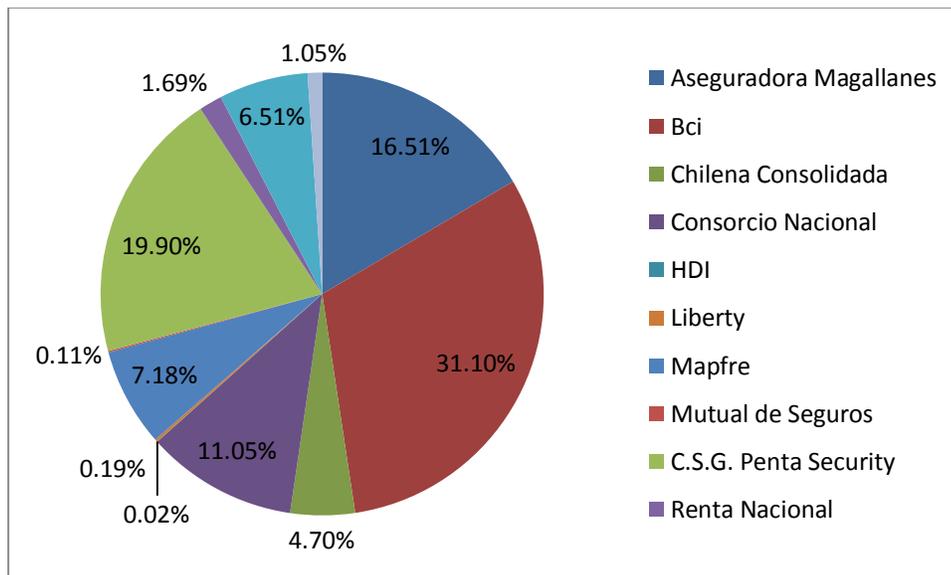


Figura 7: Numero de seguros contratados. Fuente: SVS

- Prima total:** La prima corresponde al valor del seguro (que en este caso se paga sola una vez) y representa los ingresos para las compañías de seguro. En el año 2012 la prima total en el mercado fue de 52.208.245 M\$ y su distribución se aprecia en el siguiente gráfico.

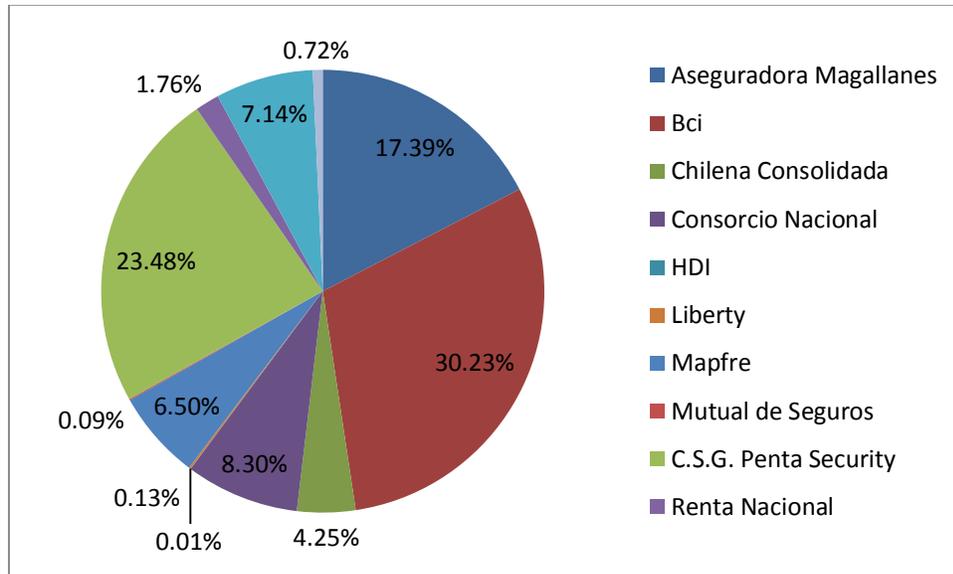


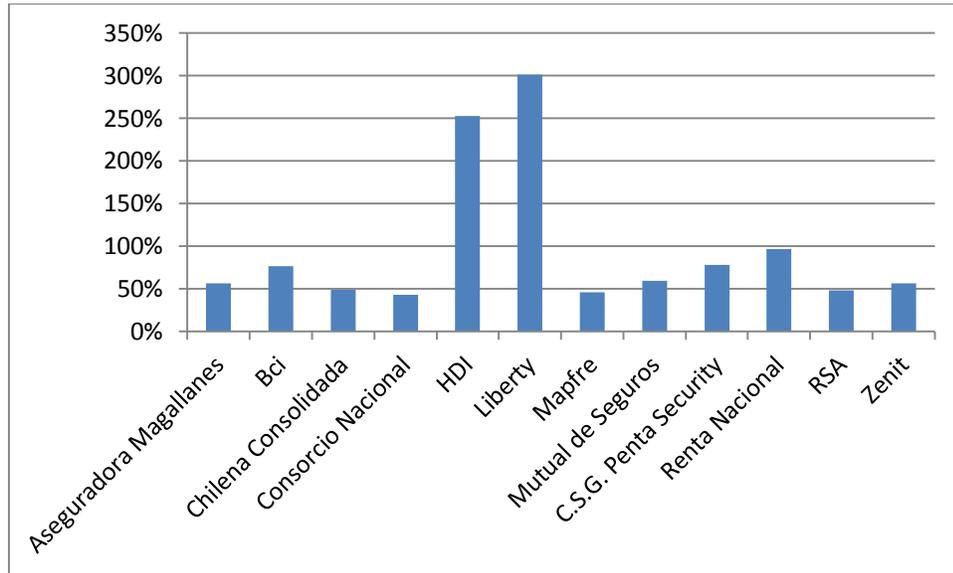
Figura 8: Prima Total. Fuente: SVS

Finalmente se puede calcular la siniestralidad comercial de cada empresa. Este indicador es muy utilizado en la industria y expresa que porcentaje de la prima total se utilizó para pagar siniestros. Mientras menor sea este indicador expresa un resultado a priori mejor (aunque no se están considerando otros costos y comisiones que podrían provocar malos resultados a pesar de obtener una buena siniestralidad). En el caso de que este indicador sea mayor que 1 se entiende que la compañía pago más dinero del que ingreso por concepto de prima.

La fórmula de la siniestralidad comercial (en su forma general) se puede expresar de la siguiente forma:

$$Siniestralidad: \frac{(Siniestros Totales + OYNR)}{Prima Total}$$

Realizando el cálculo anterior, se obtiene la siguiente figura, que muestra la siniestralidad por empresas. Se puede apreciar que Consortio tiene la mejor siniestralidad con un 43%.



**Figura 9: Siniestralidad. Fuente: SVS**

El resultado anterior nos da un vistazo de la estrategia de Consorcio, ya que, se hace visible la preferencia por ciertos segmentos (los más rentables) y el intento de capturar más valor de carteras de baja siniestralidad que pueden generar márgenes considerables.

### 1.3. Objetivo General

Aumento del beneficio y marketshare del seguro obligatorio de accidentes personales (SOAP) mediante la automatización de procesos segmentación de clientes y pricing.

### 1.4. Alcances

- El proyecto se limita al estudio de los últimos 5 años de historia para el producto SOAP
- Para el cálculo de OYNR se implementara una versión modificada de la metodología “Chain-Leader”.
- Para la segmentación se utilizan los algoritmos de K-Means y conglomerado bietapico, mientras que para la clasificación se utilizan redes neuronales y árboles de decisión.
- La automatización de los procesos son soportadas por rutinas de ETL.
- Se presenta un prototipo del sistema TI. Con respecto al proyecto no se considera la implementación final de los sistemas diseñados.
- El trabajo contempla los estudios y evaluaciones correspondientes a la metodología de ingeniera de negocios.

## Capítulo 2

### Marco Conceptual

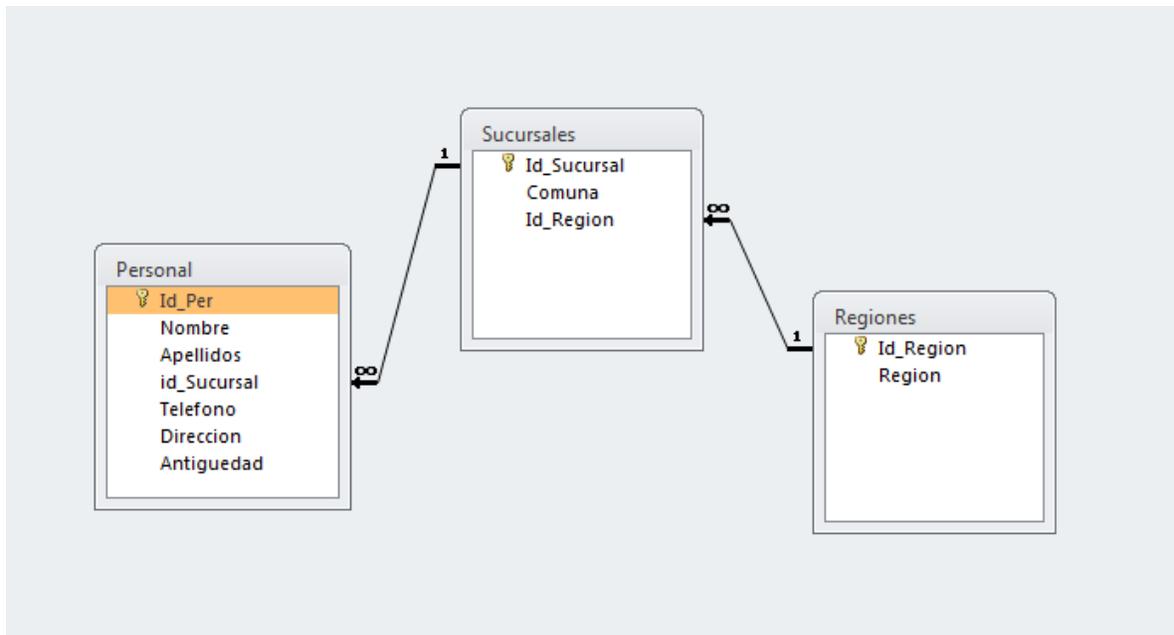
#### 2.1. Marco Teórico

En el presente trabajo se utilizan ciertas herramientas que son fundamentales para la creación e implementación de los sistemas de apoyo TI, entre ellas destacan las bases de datos relacionales y los algoritmos de data mining.

##### 2.1.1. Bases de datos Relacional

Edgar Frank Codd definió las bases del modelo relacional a finales de los 60. En las bases de Codd se definían los objetivos de este modelo [12]:

- **Independencia física:** La forma de almacenar los datos, no debe influir en su manipulación lógica.
- **Independencia lógica:** Las aplicaciones que utilizan la base de datos no deben ser modificadas porque se modifiquen elementos de la base de datos.
- **Flexibilidad:** La base de datos ofrece fácilmente distintas vistas en función de los usuarios y aplicaciones.
- **Uniformidad:** Las estructuras lógicas siempre tienen una única forma conceptual (las tablas).
- **Sencillez.**



**Figura 10: Ejemplo modelo Relacional. Fuente: Elaboración Propia.**

En forma concisa una base de datos relacional es una base de datos que almacena información sobre los de datos y cómo estos se relacionan entre ellos. En la estructuración relacional, todos los datos y las relaciones se representan en tablas bidimensionales planas llamadas relaciones. [13].

El software de base de datos relacional se llama el sistema de gestión de bases de datos relacionales (RDBMS); que controla la lectura, la escritura, la modificación, y el procesamiento de la información almacenada en las bases de datos. Los datos son descritos formalmente y organizados de acuerdo a cada modelo relacional (esquema de base de datos), en concordancia con el diseño.

Cada base de datos es una colección de tablas relacionadas. Cada tabla es una representación física de una entidad o de objeto que está en un formato tabular que consiste en columnas y filas. Las columnas son los campos de un registro o los atributos de una entidad. Las filas contienen los valores o instancias de datos; estos también son llamados registros o tuplas.

Las relaciones existen tanto entre las columnas dentro de una tabla como entre las tablas. Estas relaciones tienen tres formas lógicas: uno-a-uno, uno-a-muchos o muchos-a-muchos. La mayoría de bases de datos relacionales están diseñadas de manera que sólo hay un valor por celda (una intersección de una columna y fila); en este patrón de diseño, sólo hay uno-a-uno en una tabla. Cada tabla es nombrada de acuerdo a los datos que contiene (Ej: personas o direcciones).

Para que un DBMS pueda operar de manera eficiente y precisa, debe utilizar transacciones ACID<sup>4</sup>. Parte de este procesamiento implica ser constantemente capaz de seleccionar o modificar una y sólo una fila de una tabla. Por lo tanto, la mayoría de las implementaciones físicas tienen

<sup>4</sup> ACID: Acrónimo para; Atomicity, Consistency, Isolation, Durability. Set de propiedades que garantizan que las transacciones de bases de datos se procesen de forma confiable.

asignados por el sistema, claves primarias únicas para cada tabla. Cuando una nueva fila se escribe en la tabla, el sistema genera y escribe el nuevo valor único para la clave principal (PK); esta es la clave de que el sistema utiliza principalmente para acceder a la tabla. El rendimiento del sistema se ha optimizado para las PK. Otras claves más naturales también pueden ser identificadas y definidas como claves alternativas (AK). A menudo, pueden ser necesarias varias columnas para formar una AK (esta es una razón para que una sola columna de número entero se use generalmente como la PK). Ambas PKs y AKs tienen la capacidad de identificar de forma única una fila dentro de una tabla.

Las claves primarias dentro de una base de datos se utilizan para definir las relaciones entre las tablas. Cuando un PK migra a otra tabla, se convierte en una clave foránea de la otra tabla. Cuando cada celda sólo puede contener un valor y el PK migra en una tabla de entidad regular, este patrón de diseño puede representar un uno de uno a, o una relación de uno a muchos. La mayoría de las bases de datos relacionales son diseñadas para resolver “muchos-a-muchos” mediante la creación de una tabla adicional que contiene los PKs de las otras tablas de entidad, la relación se convierte en una entidad; la tabla de resolución es nombrada de forma apropiada y con frecuencia se le asigna su propia PK mientras que los dos FKs se combinan para formar un AK. La migración de PK a otras tablas es la segunda razón en importancia por la que los números enteros asignados por el sistema se utilizan normalmente como PK. [14]

La mayor parte de la programación dentro de un RDBMS se lleva a cabo utilizando procedimientos almacenados (SP). A menudo, los procedimientos pueden utilizarse para reducir en gran medida la cantidad de información transferida dentro y fuera de un sistema. Para mayor seguridad, el diseño del sistema también podrá conceder el acceso sólo a los procedimientos almacenados y no directamente a las tablas. Procedimientos almacenados Fundamentales contienen la lógica necesaria para insertar nuevos datos y actualizar los datos existentes. Procedimientos más complejos pueden ser escritos para implementar reglas y lógicas adicionales relacionadas con el procesamiento o la selección de los datos.

Las bases de datos relacionales tienden a ser significativamente normalizadas o des-normalizadas. Bases de datos normalizadas (bases de datos operacionales) tienden a tener una mejor integridad de los datos, reducción de espacio de almacenamiento, y transmisiones más rápidas. Las bases de datos son “des-normalizadas” para mejorar el rendimiento, a menudo para aplicaciones de toma de decisiones (análisis de tendencia) e informes. Bases de “datos des-normalizadas” pueden formar Datamarts y Data Warehouse. La mayoría de los Data Warehouse tienen una base de datos operativa que se utiliza para producir los Datamarts [12]. [13].

#### **2.1.1.1. Data Warehouse**

Un Data Warehouse es una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), que ayuda a la toma de decisiones en la entidad en la que se utiliza. Se trata, sobre todo, de un expediente completo de una organización, más allá de la información transaccional y operacional, almacenado en una base de datos diseñada para favorecer el análisis y la divulgación eficiente de datos. El almacenamiento de los datos no debe usarse con datos de uso actual. Los almacenes de datos contienen a menudo grandes cantidades de información que se subdividen a veces en unidades lógicas más pequeñas dependiendo del subsistema de la entidad del que procedan o para el que sean necesario [14].

Según definió Bill Inmon, un Data Warehouse se caracteriza por ser [15]:

- **Integrado:** los datos almacenados en el Data Warehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.
- **Temático:** sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del Data Warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.
- **Histórico:** el tiempo es parte implícita de la información contenida en un Data Warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el Data Warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el Data Warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.
- **No volátil:** el almacén de información de un Data Warehouse existe para ser leído, y no modificado. La información es por tanto permanente, significando la actualización del Data Warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

En un Data Warehouse lo que se quiere es contener datos que son necesarios o útiles para una organización, es decir, que se utiliza como un repositorio de datos para posteriormente transformarlos en información útil para el usuario. Un Data Warehouse debe entregar la información correcta a la gente indicada en el momento óptimo y en el formato adecuado. El almacén de datos da respuesta a las necesidades de usuarios expertos, utilizando Sistemas de Soporte a Decisiones (DSS), Sistemas de información ejecutiva (EIS) o herramientas para hacer consultas o informes. Los usuarios finales pueden hacer fácilmente consultas sobre sus almacenes de datos sin tocar o afectar la operación del sistema.

En el funcionamiento de un almacén de datos son muy importantes las siguientes ideas [16]:

- Integración de los datos provenientes de bases de datos distribuidas por las diferentes unidades de la organización y que con frecuencia tendrán diferentes estructuras (fuentes heterogéneas). Se debe facilitar una descripción global y un análisis comprensivo de toda la organización en el almacén de datos.
- Separación de los datos usados en operaciones diarias de los datos usados en el almacén de datos para los propósitos de divulgación, de ayuda en la toma de decisiones, para el análisis y para operaciones de control. Ambos tipos de datos no deben coincidir en la misma base de datos, ya que obedecen a objetivos muy distintos y podrían entorpecerse entre sí.

Es práctica común normalizar los datos antes de combinarlos en el almacén de datos mediante herramientas de extracción, transformación y carga (ETL). Estas herramientas leen los datos primarios (a menudo bases de datos OLTP de un negocio), realizan el proceso de transformación al almacén de datos (filtración, adaptación, cambios de formato, etc.) y escriben en el almacén.

Finalmente es usual la creación de Datamarts (subconjuntos de datos con el propósito de ayudar a que un área específica dentro del negocio) a partir del Data Warehouse (aunque en algunas metodologías se realiza el proceso inverso y el Data Warehouse es construido a partir de los Datamarts).

En la siguiente imagen se puede apreciar un esquema tipo sobre un Datawarehouse y sus componentes principales.

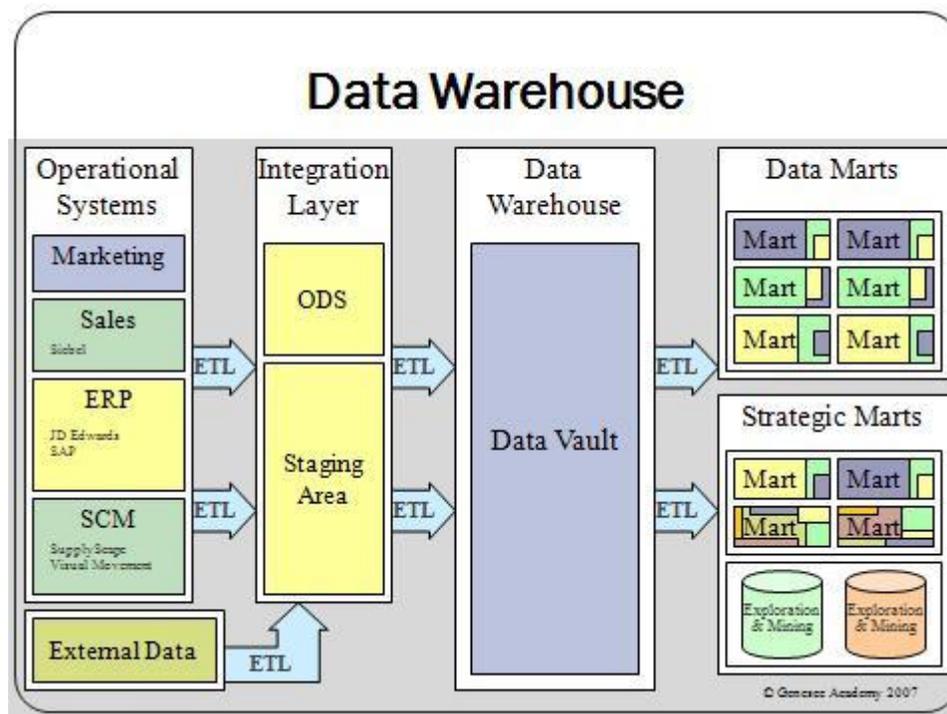


Figura 11: Esquema Data Warehouse [24]

### 2.1.2. Data Mining

El Data Mining (o minería de datos) es un subcampo interdisciplinario de las ciencias de la computación y se puede definir como el proceso de cálculo encargado de descubrir patrones en grandes conjuntos de datos utilizando herramientas propias de la inteligencia artificial, aprendizaje automático, estadísticas y sistemas de bases de datos. El objetivo general del proceso de Data Mining consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior (toma de decisiones de negocio) [17]. Además de la etapa de análisis en propiamente tal, se tratan aspectos de gestión de datos de aspectos, pre-procesamiento, consideraciones de modelo e inferencia, métricas de interés, consideraciones complejas, post-procesamiento de las estructuras descubiertas, la visualización y actualización en línea [17].

La tarea principal del Data Mining es el análisis automático o semi-automático de grandes cantidades de datos para extraer patrones interesantes hasta ahora desconocidos, como los grupos de registros de datos (análisis de conglomerados), registros inusuales (detección de anomalías) y dependencias (reglas de asociación). Esto generalmente implica el uso de técnicas de bases de datos como índices espaciales. Estos patrones pueden ser vistos como una especie de resumen de los datos de entrada, y se pueden utilizar en otros análisis o, por ejemplo, en el aprendizaje automático y análisis predictivo. Por ejemplo, la etapa de minería de datos puede identificar varios grupos en los datos, que luego se pueden utilizar para obtener resultados más precisos de predicción por un sistema de soporte de decisiones. Ni la recopilación de datos, preparación de datos, ni la interpretación de resultados y presentación de informes son parte de la etapa de extracción de datos, pero que pertenecen al proceso general KDD<sup>5</sup> como pasos adicionales.

En forma de resumen se puede decir que un proceso típico de minería de datos consta de los siguientes pasos generales [18]:

- Selección del conjunto de datos, tanto en lo que se refiere a las variables objetivo (aquellas que se quiere predecir, calcular o inferir), como a las variables independientes (las que sirven para hacer el cálculo o proceso), como posiblemente al muestreo de los registros disponibles.
- Análisis de las propiedades de los datos, en especial los histogramas, diagramas de dispersión, presencia de valores atípicos y ausencia de datos (valores nulos).
- Transformación del conjunto de datos de entrada, se realizará de diversas formas en función del análisis previo, con el objetivo de prepararlo para aplicar la técnica de minería de datos que mejor se adapte a los datos y al problema, a este paso también se le conoce como pre-procesamiento de los datos.
- Seleccionar y aplicar la técnica de minería de datos, se construye el modelo predictivo, de clasificación o segmentación.
- Extracción de conocimiento, mediante una técnica de minería de datos, se obtiene un modelo de conocimiento, que representa patrones de comportamiento observados en los valores de las variables del problema o relaciones de asociación entre dichas variables. También pueden usarse varias técnicas a la vez para generar distintos modelos, aunque generalmente cada técnica obliga a un pre-procesado diferente de los datos.
- Interpretación y evaluación de datos, una vez obtenido el modelo, se debe proceder a su validación comprobando que las conclusiones que arroja son válidas y suficientemente satisfactorias. En el caso de haber obtenido varios modelos mediante el uso de distintas técnicas, se deben comparar los modelos en busca de aquel que se ajuste mejor al problema. Si ninguno de los modelos alcanza los resultados esperados, debe alterarse alguno de los pasos anteriores para generar nuevos modelos.

---

<sup>5</sup> KDD: Acronimo para Knowledge discovery in databases.

En la siguiente figura se puede apreciar cómo funciona el proceso de Data Mining es un caso genérico:

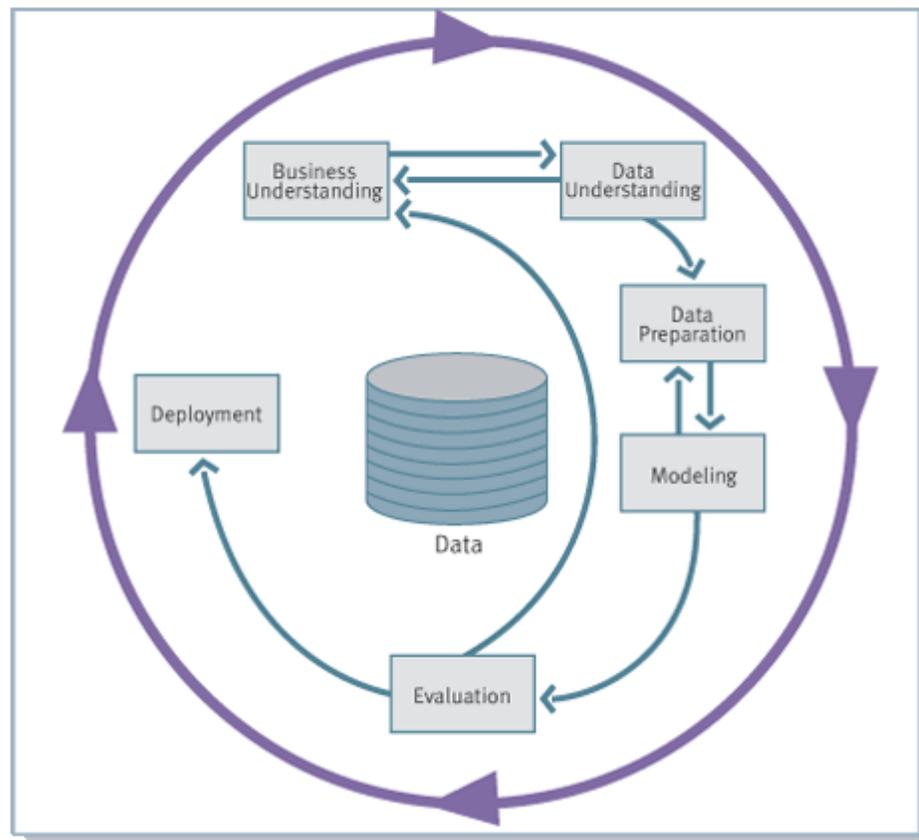


Figura 12: Esquema Data Mining [19]

Dentro del presente proyecto se utilizarán principalmente dos herramientas de Data Mining, las cuales se explican a continuación.

### 2.1.2.1. Clustering

El análisis de agrupamiento o clustering es la tarea de agrupar un conjunto de objetos de tal manera que los objetos en el mismo grupo (llamado cluster) sean más similares (en un sentido u otro) entre sí que a los de otros grupos (clusters). Es una de las tareas principales de exploración del Data Mining, y una técnica común para estadísticos de análisis de datos, se utiliza en muchos campos, incluyendo el aprendizaje automático, reconocimiento de patrones, análisis de imágenes, recuperación de la información hasta la bioinformática. [18].

El análisis de conglomerados no es un algoritmo específico, sino el problema general a resolver. Se puede lograr por diversos algoritmos que difieren significativamente en su noción de lo que constituye un clúster y cómo construirlos de manera eficiente. Las nociones populares de grupos incluyen grupos con pequeñas distancias entre los miembros del clúster, áreas densas del espacio de datos, intervalos o distribuciones estadísticas particulares. El clustering, por lo tanto se puede formular como un problema de optimización multiobjetivo. Los ajustes del algoritmo de

conglomeración y parámetros apropiados (incluidos los valores tales como la función de distancia a utilizar, un umbral de densidad o el número de conglomerados esperados) dependen del conjunto de datos individuales y el uso de los resultados previstos. El análisis de clusters, como tal, no es una tarea automática, sino un proceso iterativo de descubrimiento de conocimiento o la optimización multiobjetivo interactiva que implica prueba y error. A menudo será necesario modificar los parámetros de preprocesamiento y modelo de datos hasta que el resultado alcanza las propiedades deseadas. [20]

Asimismo del término clustering cuenta con una serie de términos con significados similares, incluyendo clasificación automática, taxonomía numérica, botryology (del griego βότρυς "uva") y el análisis tipológico. Las diferencias sutiles son a menudo en el uso de los resultados: mientras que en el Data Mining, los grupos resultantes son el tema de interés, en la clasificación automática el poder discriminativo resultante es de mayor relevancia. Esto conduce a menudo a malentendidos entre los investigadores procedentes de los campos del data mining y aprendizaje automático, ya que utilizan los mismos términos y con frecuencia los mismos algoritmos, pero tienen diferentes objetivos.

Algunos de los algoritmos de Clustering más conocidos y utilizados por los profesionales del área son los siguientes: [20]

- K-means clustering
- K-medians clustering
- K-medoids
- K-nearest neighbors algorithm
- Self-organizing map
- Nearest-neighbor chain algorithm
- Affinity propagation
- Hierarchical Clustering
- Fuzzy clustering

#### **2.1.2.2. Clasificación**

En el aprendizaje automático y las estadísticas, la clasificación es el problema de identificar a cuál de un conjunto de categorías (sub-poblaciones) una nueva observación pertenece, sobre la base de un conjunto de datos de entrenamiento que contiene las observaciones (o instancias) cuya pertenencia categoría se conoce. [18]

En la terminología de aprendizaje automático, la clasificación se considera una instancia de aprendizaje supervisado, es decir, aprender sobre un conjunto de entrenamiento de observaciones correctamente identificadas. El proceso correspondiente sin supervisión corresponde al Clustering

(sección anterior), y consiste en agrupar los datos en categorías sobre la base de un cierto grado de similitud inherente o distancia. [20]

A menudo, las observaciones individuales se analizan en un conjunto de propiedades cuantificables, variables explicativas, características, etc. Estas propiedades pueden ser categóricas, ordinales, de valor entero o de valor real. Otros clasificadores trabajan mediante la comparación de observaciones a las observaciones anteriores por medio de una similitud o función de distancia.

Un algoritmo que implementa la clasificación, especialmente en una aplicación concreta, es conocido como un clasificador. El término "clasificador" a veces también se refiere a la función matemática implementada por un algoritmo de clasificación, que mapea datos de entrada a una categoría. [21]

La terminología intra-campos es bastante variada. En las estadísticas, en la que la clasificación se hace a menudo con la regresión logística o un procedimiento similar, las propiedades de las observaciones se denominan variables explicativas (o variables independientes, regresores, etc), y las categorías que se predijo que se conoce como los resultados, que se consideran a ser posibles valores de la variable dependiente . En el aprendizaje automático, las observaciones son a menudo conocidas como los casos, las variables explicativas son denominadas características (agrupados en un vector de características), y las posibles categorías para predecir son las clases. También existe cierta controversia sobre si los métodos de clasificación que no implican un modelo estadístico se pueden considerar "estadística". Otros campos pueden usar terminología diferente: por ejemplo, en la ecología de comunidades , el término "clasificación" se refiere normalmente a análisis de conglomerados , es decir, un tipo de aprendizaje no supervisado, en lugar del aprendizaje supervisado recientemente descrito.

Ejemplos de algoritmos de clasificación incluyen: [21]

- Clasificadores lineales
  - Discriminante lineal de Fisher
  - La regresión logística
  - Naive Bayes Classifier

- Perceptron
- Support Vector Machines (SVM)
  - Least Squares Support Vector Machine
- Clasificadores cuadráticas
- Estimación Kernel
  - k-vecinos más cercanos
- Impulsar (meta-algoritmo)
- Los árboles de decisión (decision trees)
  - Random forest
- Las redes neuronales (Artificial neural network)

### 2.1.2.3. Técnicas de Evaluación

Conforme verificar que algoritmo o parámetros utilizar es necesario utilizar técnicas para evaluar y comparar distintas opciones para de esta forma poder tomar las mejores decisiones posibles.

En el presente trabajo se utilizan principalmente tres indicadores; El coeficiente de Gini, el índice Davies–Bouldin y la suma de cuadrados, los cuales son descritos brevemente a continuación:

- Coeficiente de Gini: El coeficiente de Gini (también conocido como el índice de Gini o coeficiente de Gini) es una medida de dispersión estadística. Mide la desigualdad entre los valores de una distribución de frecuencias. Un bajo coeficiente de Gini indica una distribución más equitativa, 0 corresponde a completa igualdad, mientras más altos coeficientes de Gini indican una distribución más desigual, con 1 completa desigualdad. Esto puede ser aplicada a algoritmos de Clustering para verificar cuan dispersos los datos están en los distintos cluster.
- Índice Davies–Bouldin: El índice de Davies-Bouldin (DBI) es una métrica para la evaluación de algoritmos de clustering. Se trata de un esquema de evaluación interna, donde la validación de lo bien que la agrupación se ha realizado depende de cantidades y características inherentes al set de datos. El índice es simétrico y representa la relación ente la distancia intra-cluster e inter cluster, en donde valores menores presentan mejores resultados.

- Suma de Cuadrados: La suma de cuadrados hace referencia a la medida promedio de varianza de cada cluster, mientras menor sea este indicador, la varianza ha sido minimizada en los cluster y por ende se ha realizado un mejor Clustering.

## 2.2. Marco Metodológico

### 2.2.1. Ingeniería de Negocios

La Ingeniería de Negocios es la disciplina que provee los fundamentos y la metodología que permiten diseñar una empresa, incluyendo su Arquitectura Empresarial, de la cual es parte la arquitectura de procesos, y el detalle de todos los procesos necesarios para que la empresa sea competitiva.

En la siguiente figura se puede apreciar un resumen de la metodología:

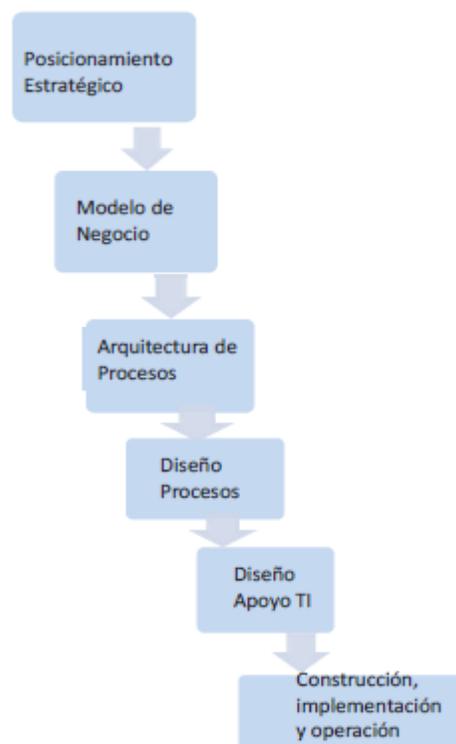


Figura 13: Metodología Ingeniería de Negocios [6].

La ingeniería de negocios es un elemento clave a la hora de diseñar los procesos de negocios de una organización. Uno de los aspectos más importantes en el diseño de procesos es la existencia de varios niveles de diseño, los cuales deben ser vistos de manera coordinada. Los niveles definidos son los siguientes [6]:

- Modelo de negocio y arquitectura de procesos y su relación con la estrategia y la estructura organizacional, también llamada arquitectura empresarial.
- Diseño de procesos que implementan el modelo, con énfasis en el negocio.
- Diseño y construcción de las aplicaciones TI de apoyo a los procesos

Conforme tener un proceso de ingeniería de negocios exitoso y efectivo, es necesario seguir la secuencia presentada en la figura 13, la cual de forma desagrada se queda explicitada a continuación.

- **Planteamiento Estratégico:** punto de partida de la metodología que busca generar las líneas que sigue la empresa con respecto a su posicionamiento estratégico. En esta etapa se detalla el posicionamiento estratégico de la empresa de acuerdo a los trabajos de M. Porter y A. Hax. Asimismo, se busca generar un mapa estratégico donde se relacione la visión y misión de la empresa con las distintas perspectivas del BSC hasta llegar a la definición de los procesos internos de nivel más bajo.
- **Definición del Modelo de Negocio:** se describe cómo se pretende llevar a cabo el posicionamiento estratégico propuesto por la empresa generando valor al cliente a través de una propuesta de valor. Para el desarrollo de este punto se utilizará en enfoque planteando por Johnson, Christensen, y Kagermann.
- **Diseño de la Arquitectura de Procesos:** se realiza el diseño de la arquitectura de Macroprocesos (capítulo 4) a partir del modelo de negocios. Es importante recalcar que lo más relevante que se propone en esta arquitectura de procesos son explicitar las relaciones que hay entre estos. Este diseño de Macroprocesos se realiza utilizando herramientas computacionales que permitan modelar bajo la metodología IDEF0.
- **Diseño Detallado de los Procesos:** una vez detallada la arquitectura de Macroprocesos, se realiza un diseño más minucioso utilizando como punto de partida los Patrones de Procesos de Negocios. Para este diseño también se utiliza una herramienta computacional que permita modelar bajo la metodología BPMN
- **Diseño de la(s) Aplicación(es) de Apoyo:** generado a partir de los diseños de procesos en BPMN del punto anterior. Para cada modelo BPMN se determina el apoyo requerido en términos computaciones y luego se procede a diseñar utilizando la metodología UML y una programación orientada a objetos.
- **Construcción e Implementación de la Solución:** Por último, ya con toda la información necesaria y detallada en los pasos anteriores, se finaliza con la implementación de un sistema de apoyo. Aquí se construyen la o las aplicaciones necesarias para desarrollar de manera eficaz, los procesos diseñados que apoyan al modelo de negocios propuesto.

## Capítulo 3

### Análisis Estratégico

#### 3.1. Misión y Visión

**Visión:**

Ser líder en la industria aseguradora y un actor relevante en ahorro y créditos, generando una rentabilidad adecuada en todas las líneas de negocio.

**Misión:**

Ser una organización que ofrece una variedad de servicios financieros de excelencia, para permitir a los clientes satisfacer sus necesidades de prosperidad familiar y seguridad patrimonial.

Crear valor para los accionistas, construyendo relaciones de confianza y cumpliendo un rol de buen ciudadano empresarial

Ofrecer las condiciones para que las personas pertenecientes a la organización desarrollen todo su potencial, en un ambiente laboral desafiante y de respeto a los valores profesados.

#### 3.2. Mapa estratégico

El siguiente mapa estratégico muestra los principales objetivos de la compañía, haciendo la distinción entre cuatro perspectivas distintas; financiera y de los dueños, del cliente, de procesos y de aprendizaje. Como se podría esperar la perspectiva financiera es bastante común para las grandes empresas, enfocándose en rentabilizar la cartera de clientes, maximizar los retornos y controlar los riesgos.

De la perspectiva del negocio se puede extraer la propuesta de valor de la empresa, en donde se explicita generar relaciones permanentes con los clientes, brindar servicios financieros de calidad y tener distintas propuestas por segmentos.

En la perspectiva de procesos es donde se vuelve más presente la gerencia técnica (donde se desarrolla el proyecto), ya que, se tiene por objetivos aumentar la eficiencia de los proyectos, gestionar y administrar calidad de datos y realizar gestión del conocimiento.

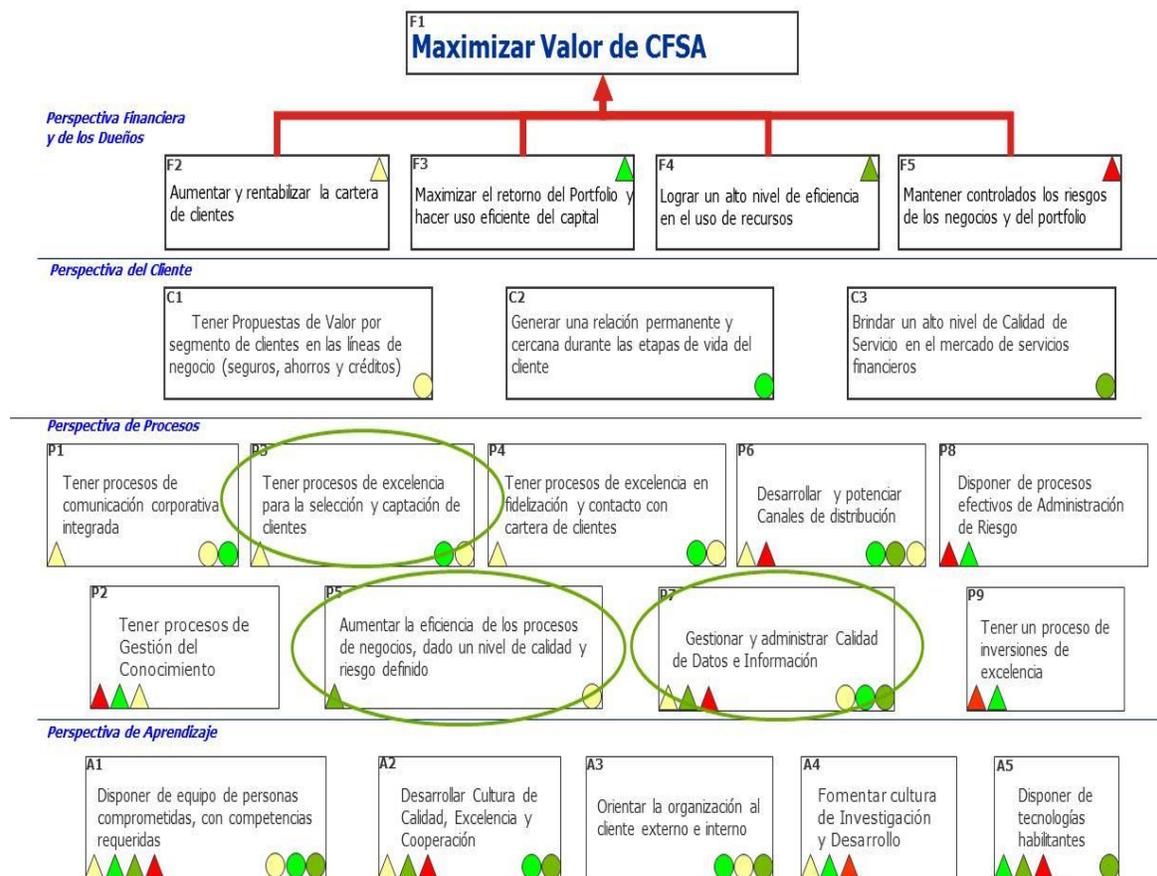


Figura 14: Mapa Estratégico CFSA [2].

Dentro del BSC se marcaron tres procesos que son fundamentales (y están alineados) con el proyecto propuesto, los cuales se detallaran a continuación

- **Tener procesos de excelencia para la selección y captación de clientes:** El proyecto genera los análisis necesarios para poder seleccionar de mejor forma a los clientes y por consiguiente crear nuevas y mejores formas de captarlos.
- **Aumentar la eficiencia de los procesos de negocios, dado un nivel de calidad y riesgo definido:** Aquí se mezclan las partes del proyecto, ya que, la automatización aumente claramente la eficacia de los procesos de negocio, pero la optimización ayuda a cuantificar y controlar mejor la calidad y el riesgo.
- **Gestionar y administrar Calidad de Datos e Información:** Parte integral del proyecto y de los sistemas TI de apoyo.

También se encuentra una BSC propia de la gerencia técnica, en donde se explican otros procesos (muy relacionados a los anteriores). Se explicita la necesidad de una tarificación óptima y mejorar el análisis y modelamiento, procesos fuertemente relacionados al proyecto.

Además se debe hacer notar la idea de la gestión de conocimiento, ya que, de poder implementarse de buena forma el proyecto, este podría replicarse para otros (sino todos) los productos de la compañía, logrando un beneficio global considerable.

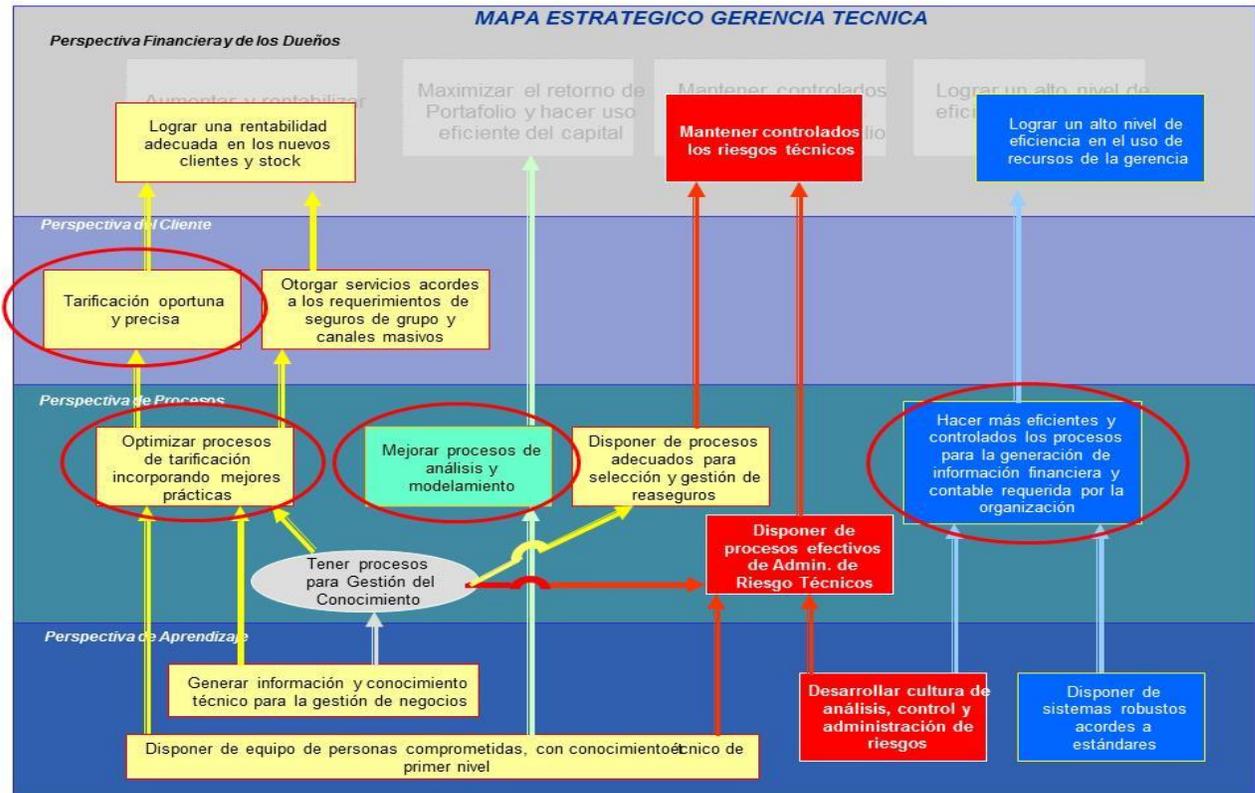


Figura 15: Mapa Estratégico Gerencia Técnica [2].

### 3.3. Modelo Delta

El Modelo Delta es un marco estratégico que sitúa al cliente al centro de la gestión. Examina las opciones primarias disponibles para establecer una vinculación con el cliente y propone como ligar la estrategia y la ejecución a través del alineamiento adaptativo en los procesos [7].

En este modelo se tienen tres posicionamientos estéricos a los que las empresas pueden aspirar, estos son: Solución Integral para los clientes, Consolidación del Sistema (Lock-In Sistémico) y Mejor producto. Dada las características de la empresa se concluye que sigue una estrategia Solución Integral para los clientes, dado que Consorcio busca crear productos personalizados basados en el conocimiento sobre sus clientes, y mantener relaciones a largo plazo donde pueda desarrollar proposiciones de mayor valor para los mismos. En general Consorcio busca ofrecer un conjunto más completo de productos y servicios, aumento el valor percibido por el cliente (esta idea es potenciada con la inclusión del Banco Consorcio hace unos años).

Vale destacar que esto se aplica a casi todos los productos, pero este no es el caso de SOAP, ya que al ser un producto obligatorio (y normado por el estado, mediante la SVS) no se puede crear

una relación más íntima o de largo plazo con los clientes, por lo que se utiliza un enfoque de mejor producto principalmente basado en el menor precio y en algunas prestaciones a los clientes como la facilidad en la compra y el reclamo de los siniestros mediante el canal web.

A continuación se presenta el llamado triángulo de HAX, en donde se explicita como afectara el proyecto a cada posicionamiento y donde se entiende porque la estrategia de la empresa (y por ende el marco en donde se erige el proyecto) tiende hacia el mejor producto y en el caso del proyecto se centra en la eficiencia operacional.

**Proyecto:**

- Disminuir costos área técnica
- Optimizar proceso de tarificación
- Desarrollo de capacidad tecnológica aplicable a otros productos
- Mayor Control
- Precios por segmento



**Figura 16: Modelo de Hax “Mejor Producto” [7].**

### 3.4. Modelo de Negocios

El modelo de negocios habla de cómo una empresa genera sus ingresos especificando en qué lugar se encuentra dentro de la cadena de valor. Una compañía de seguros entrega protección a sus asegurados, realizando una transferencia de riesgos hacia la empresa y en el caso de Consorcio ofreciendo además otra gama de servicios desde el ahorro hasta la inversión.

En la siguiente figura se puede apreciar la propuesta de valor de la empresa, explicitando los recursos claves, procesos claves y la fórmula para generar utilidades. Este modelo de negocio es el que promueve el presente proyecto, ya que, para poder cumplir a cabalidad los principios expuestos por este modelo es que se vuelve necesario la aplicación de la metodología de ingeniería de negocios y la implementación final de los sistemas TI diseñados.

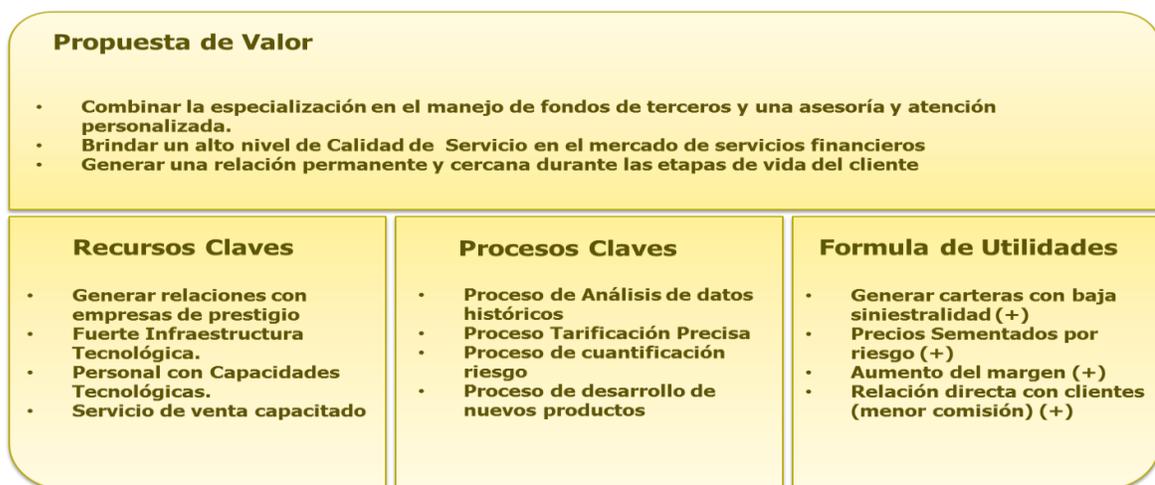
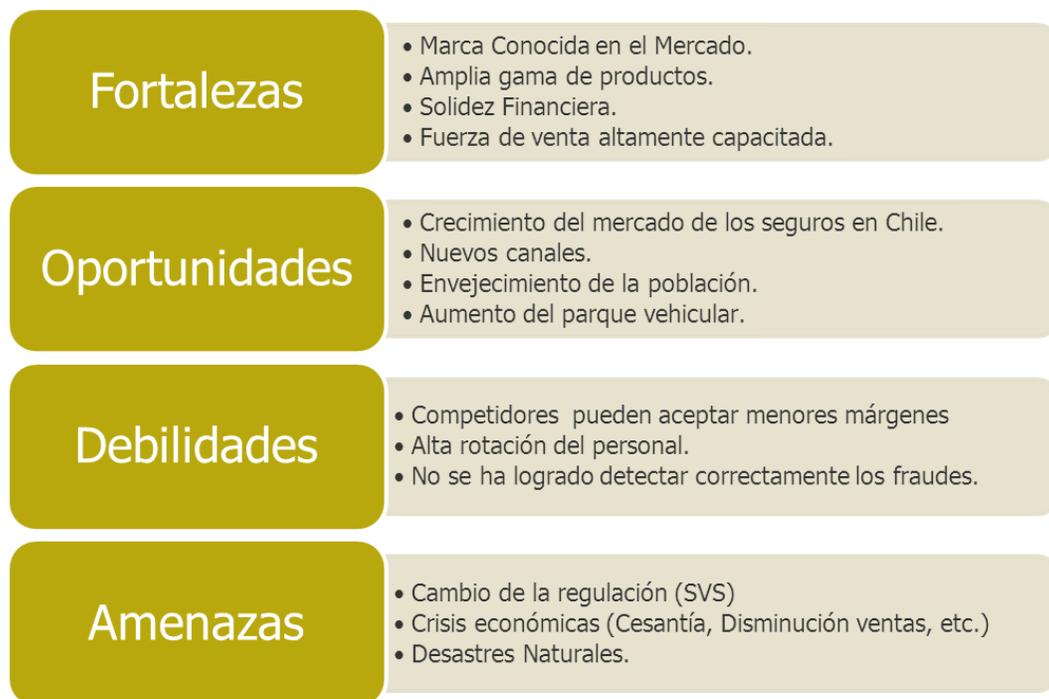


Figura 17: Modelo de negocios Consorcio. Fuente: Elaboración propia.

### 3.5. Análisis Mercado

#### 3.5.1. FODA

A continuación se puede apreciar el FODA relacionado a Consorcio.



Debido a lo anterior se desprende que Consorcio tiene grandes fortalezas y oportunidades que aprovechar. Con respecto a las amenazas, estas son compartidas por el resto de la industria mientras que las debilidades pueden ser trabajadas para reducir su impacto en la organización.

La estrategia de Consorcio debe ser analizada desde el punto de vista del FODA, por esta razón un proyecto como el presente es altamente deseable, ya que aprovecha las fortalezas fundamentales de la empresa, como la marca y la solidez financiera, y logra apropiarse, al menos en parte, de alguna de las grandes oportunidades como son los nuevos canales (venta web) y el aumento del parque vehicular del país.

### **3.5.2. Fuerzas de Porter**

Considerando el mercado de los seguros es posible analizar las 5 fuerzas de Porter [8] y graduar el nivel de cada una de estas fuerzas. Realizando el análisis se logra apreciar que el mercado de los seguros es extremadamente competitivo y de difícil diferenciación (por lo que el énfasis en la eficiencia operacional se hace crítico).

- **Rivalidad de los competidores (Alta)**
  - Gran número de competidores
  - Poca diferenciación de los productos
  - Competencia imita rápidamente productos o Procesos
  
- **Amenaza de nuevos competidores (Alta)**
  - Otras entidades financieras (bancos) entran en el negocio de los seguros.
  - Empresas extranjeras pueden entrar al mercado.
  - Empresas de retail comienzan a vender seguros (corredores).
  
- **Amenaza de productos sustitutos (Baja)**
  - No existe sustituto perfecto a los seguros.
  
- **Poder de negociación de los clientes (Alta)**
  - Bajos costos en cambio de empresa.
  - Productos poco diferenciados.
  - SVS y SERNAC intervienen en favor de los clientes.

- **Poder de negociación de los proveedores (Media)**
  - Principalmente empresas reaseguradoras.
  - Fuerte poder en negociación y renovación de contratos.
  - Existen muchos reaseguradores (además de contratos de coaseguro).

El análisis anterior hace presente lo competitivo del mercado de los seguros, así como ciertas oportunidades que pueden ser aprovechadas por las empresas de seguros. Estos factores son los que repercuten en la estrategia empresarial de Consorcio, la cual termina materializada en su misión y visión. Consorcio, así como sus competidores, deben buscar diferenciarse en una industria donde el “producto” a ser transado rápidamente se vuelve un commodity, por esta razón se debe aumentar la eficiencia de los procesos, estar más cerca del cliente (que este valore la marca y la experiencia de compra y post-venta) y encontrar pequeñas diferencias que logren aumentar los beneficios compitiendo en el mismo mercado que los demás actores de la industria reaseguradora.

En el siguiente capítulo se definirá la arquitectura de macro procesos, para luego diagramar los procesos más relevantes en donde su rediseño o mejora servirá para implementar de mejor forma la estrategia de la empresa, así como intentar obtener el conocimiento suficiente para lograr captar mayor valor de los clientes (o futuros clientes) de la empresa.

## Capítulo 4

# Arquitectura de Macro-Procesos

### 4.1. Macro-Procesos

La arquitectura diseñada se basa en los macro-procesos propuestos por Barros. Dichos macro-procesos corresponden a patrones de procesos observados en la mayoría de las empresas que presentan buenas prácticas y, por lo tanto, son extensibles a otras empresas u organizaciones, observándose que al ser implementados correctamente, dichas empresas generan eficiencia y eficacia tanto estratégica, táctica y operacionalmente.

En el ámbito de las organizaciones también es posible utilizar el término arquitectura. En efecto, la academia ha definido el concepto Arquitectura Empresarial, como la estructura de los procesos, sistemas de información, personal y unidades organizacionales de una empresa, alineados con sus objetivos fundamentales y dirección estratégica. En síntesis, una Arquitectura Empresarial puede ser entendida como el esfuerzo formal que realiza una organización para estructurar sus operaciones, sus procesos, sistemas de información, recursos y personal humano.

La Arquitectura de Macro-procesos de Oscar Barros, define cuatro grandes procesos o “macros” que debiesen estar presentes en las organizaciones. En síntesis estos son [6] [9]:

- **Macro 1 - Cadena de valor:** Macroproceso que agrupa todas las actividades que las empresas deben desarrollar para planificar, producir y entregar al cliente sus productos o servicios.
- **Macro 2 - Desarrollo de nuevas capacidades:** Macroproceso que agrupa actividades relacionadas con el estudio permanente de nuevas capacidades que la empresa debiese implementar para ser competitiva. Ejemplo de lo anterior, es la adopción de nuevas tecnologías y el desarrollo de nuevos proyectos que inciden en la cadena de valor.
- **Macro 3- Planificación estratégica:** Macroproceso que agrupa actividades relacionadas con la determinación de los lineamientos estratégicos de la organización, materializados en planes y programas de acción a ser adoptados en las operaciones de la empresa.
- **Macro 4 - Gestión de recursos habilitadores:** Macroproceso que agrupa todas aquellas actividades que dan soporte a la ejecución de los otros tres Macroprocesos. En este ámbito se encuentran la gestión de recursos humanos, infraestructura, insumos, entre otros.

A continuación se presenta la arquitectura de Macroprocesos adaptada al caso de Consorcio y al proyecto mismo. En una primera aproximación se ve el esquema general de las cuatro macros, para luego ir descendiendo a los procesos específicos en los que se enfocara el proyecto.

La macro escogida es la macro 1, o también llamada, la cadena de valor. El proyecto queda supeditado a esta macro debido a que es en esta donde se genera el valor del negocio, se gestiona producción del servicio (o producto), además de gestionar la relación con el cliente.

En el caso de Consorcio existen varias cadenas de valores, asociadas a cada tipo de servicio, en el caso de este proyecto se utiliza la cadena correspondiente al seguro en cuestión, que es “Seguros generales”.

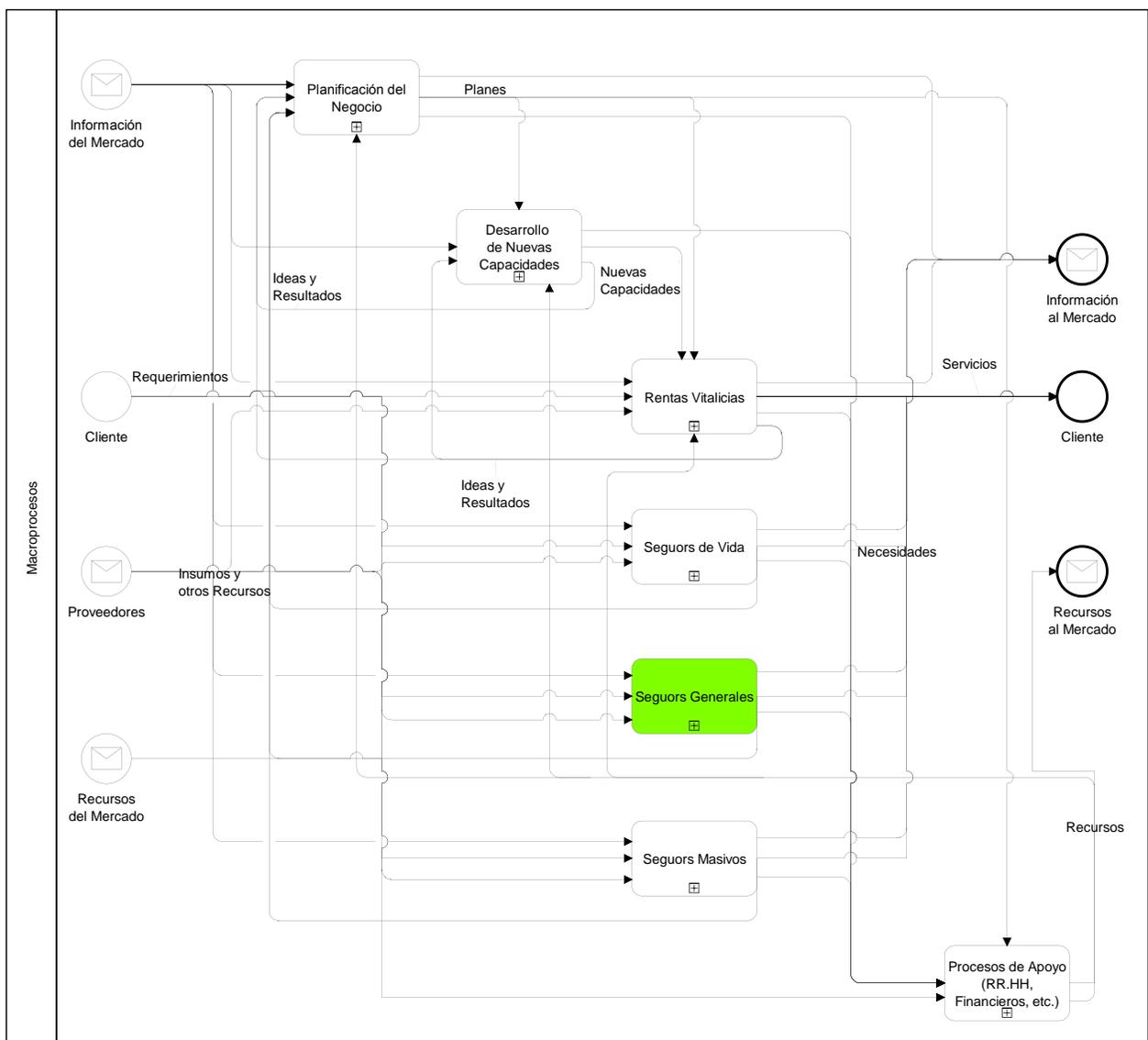


Figura 18: Macroprocesos [6].

Ya adentrándonos en la Macro 1 (ver figura siguiente) se pueden apreciar 4 procesos que la componen y se marcan en color verde en donde el proyecto se posicionara (en algunos casos necesitando de grandes cambios y en otros no).

En este caso, el proyecto tiene que estar en la primer proceso, “Administración relación con el cliente”, ya que, es aquí donde se hacen los análisis necesarios y se crean las reglas para luego vender el producto a los clientes.

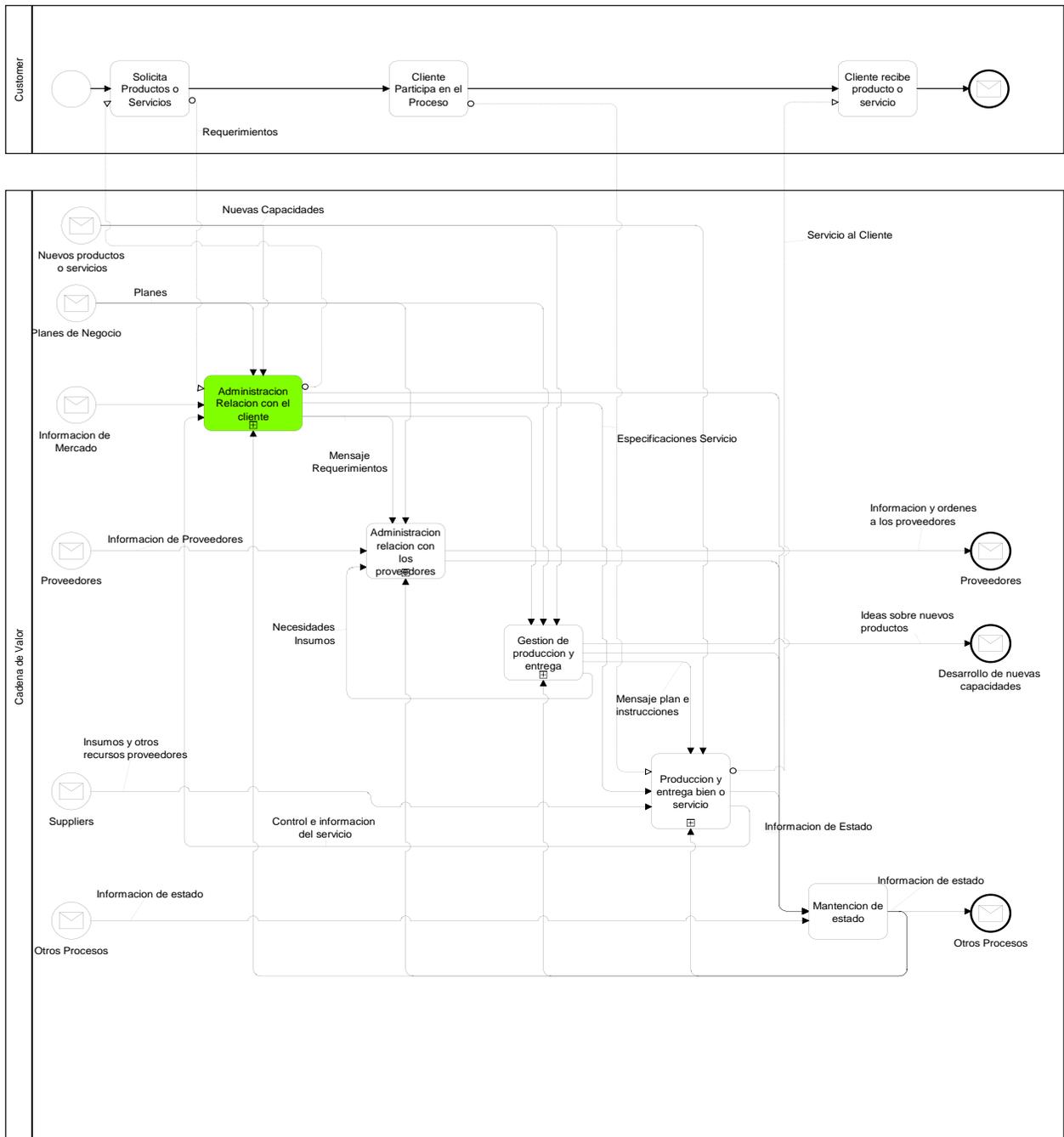


Figura 19: Macro 1, Cadena de Valor [6].

En la siguiente figura se aprecia el interior del proceso “administración relación con el cliente” y se marca nuevamente el proceso más relevante para el proyecto, en este caso marketing y análisis de mercado (ya que, se quiere segmentar el mercado y analizar su comportamiento) es aquí donde se crearan las reglas para la venta y atención de cliente (aunque el proyecto no tendrá demasiada injerencia ahí).

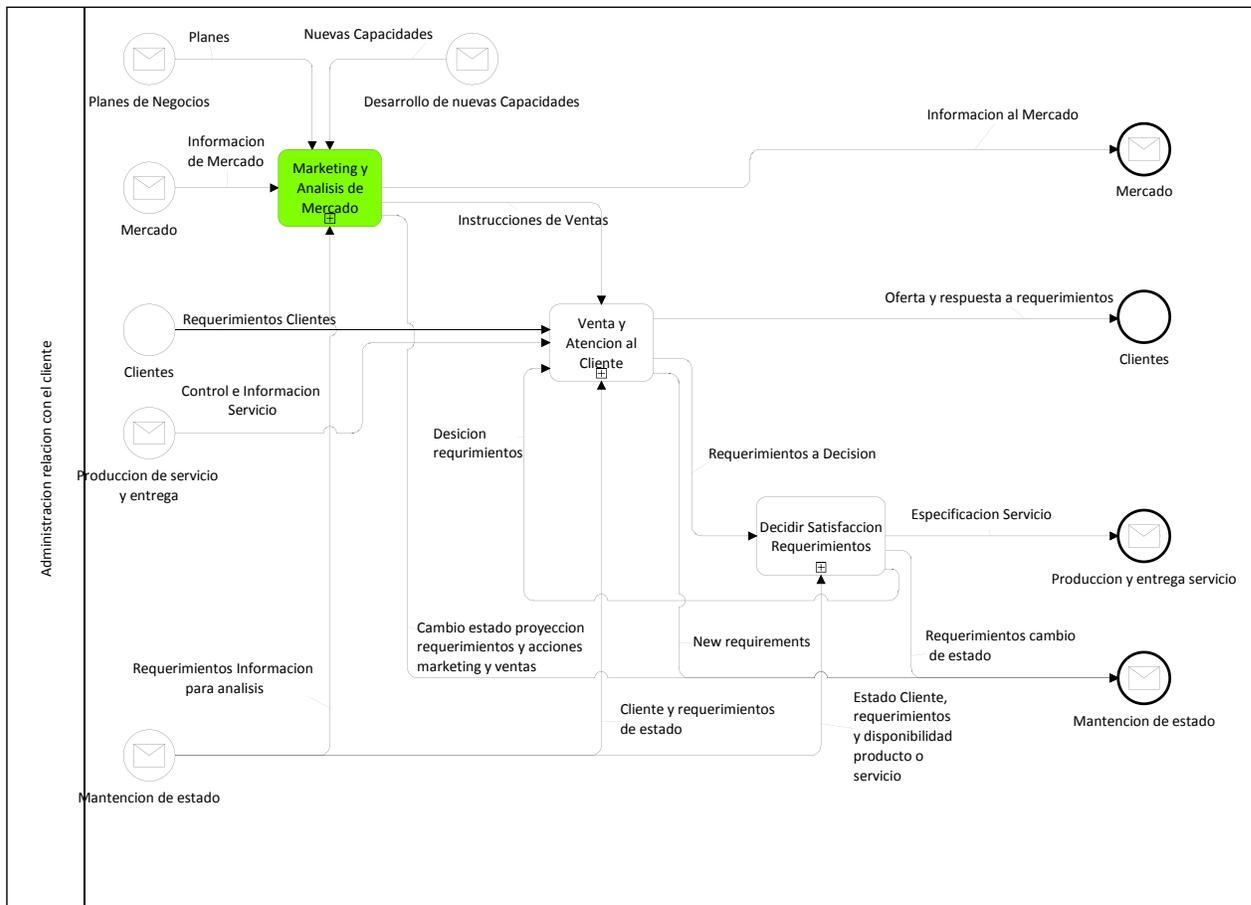


Figura 20: Administración Relación con el Cliente [6].

Dentro de marketing y análisis de mercado encontramos “Analizar comportamiento de los clientes”, que es justo el último paso antes de que el proceso de la gestión del seguro se ejecute, aquí es donde se analiza los clientes en base a la data obtenida y con este análisis se toman las decisiones pertinentes de marketing y la planificación de las ventas.

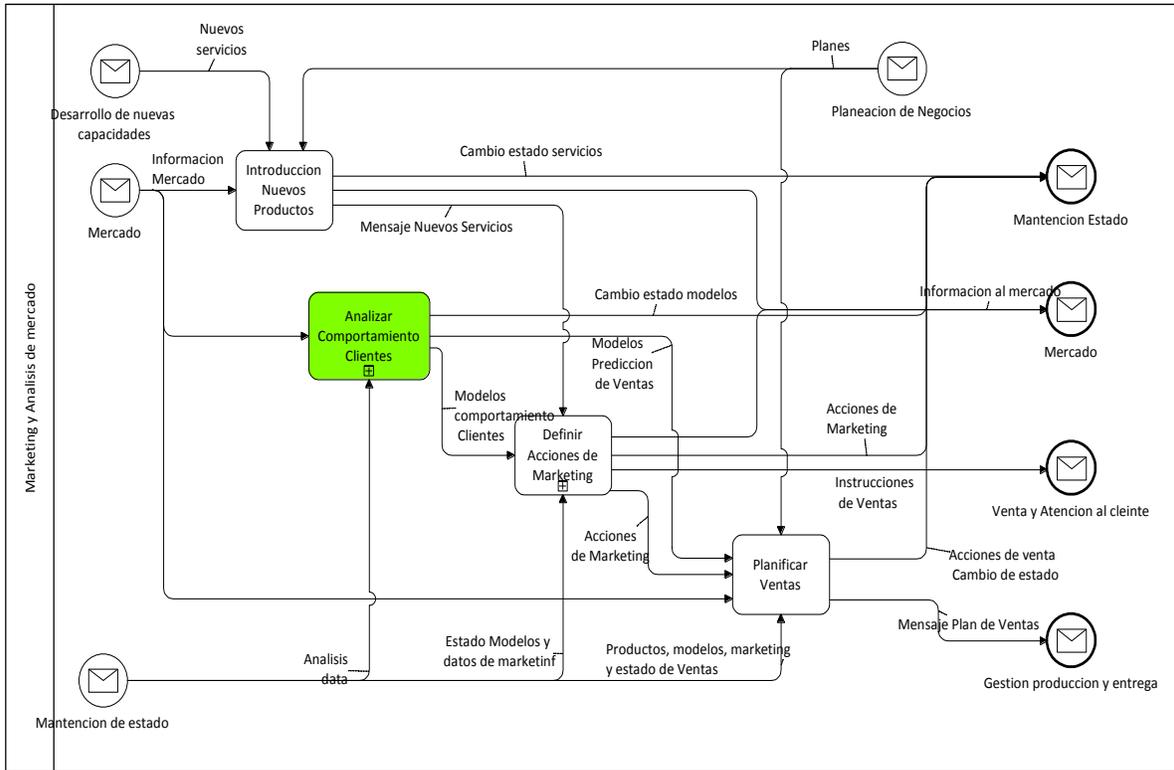


Figura 21: Marketing y análisis de mercado [6].

En el proceso de “Analizar el comportamiento de los clientes”, se encuentran 3 procesos relevantes para todo el diseño posterior, que son la carga de datos al data Warehouse, la preparación de los datos necesarios para los modelos y finalmente la ejecución de los modelos escogidos, lo que permite completar el objetivo de segmentar y encontrar los precios buscados.

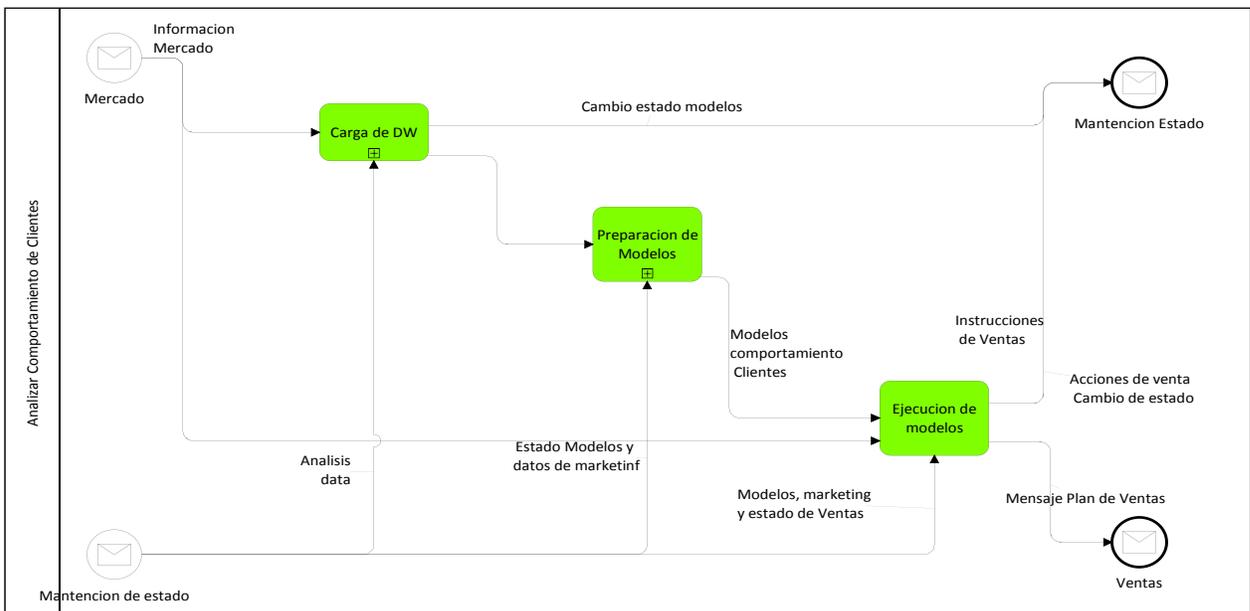


Figura 22: BP1. Fuente: Elaboración propia.

## 4.2. Variables de diseño

Conforme poder realizar el rediseño de los proceso de negocio que se desprenden del proyecto se analizan las variables de diseño relevantes [6]. Se estudian seis variables principales; Estructura Empresa Mercado, Anticipación, Coordinación, Practicas de trabajo, Integración de Procesos Conexos y Mantención Consolidada de Estado (y como están se presentan en la empresa), haciendo la comparación entre la situación actual y la propuesta por el proyecto.

### 4.2.1. Estructura Empresa Mercado

Esta es la variable de mayor impacto sobre el proceso y está presente cuando, al nivel de estrategia, modelo de negocio y de arquitectura, se ha decidido hacer cambios significativos en la estructura de negocio y los procesos y/o en las relaciones con clientes y proveedores. Estos cambios implican siempre importantes modificaciones en la estructura organizacional y, en forma subyacente, siempre debe existir el apoyo TI adecuado para permitir este cambio.

a. Estructura Empresa Mercado	Actual	Proyecto
a.1. Servicio Integral al Cliente.	No	Si (Pero el servicio integral escapa del proyecto, aunque existe)
a.2. Lock-in Sistémico.	No	No
a.3. Integración con Proveedores.	No	No
a.4. Estructura Interna: centralizada o descentralizada.	Centralizada	Operación descentralizada, Lógica de Negocio Centralizada
a.5. Toma de Decisiones: Centralizada o Descentralizada.	Descentralizada	Descentralizar decisiones con lógicas de negocio centralizadas y automatizadas.

**Tabla 1: Estructura Empresa Mercado. Fuente: Elaboración propia.**

### 4.2.2. Anticipación

Esta variable debe su existencia a que una de las necesidades más importantes en las organizaciones es la de anticiparse a los eventos futuros (en la medida de lo posible). Obviamente, tal anticipación requiere de una capacidad de predicción de tales eventos. Uno de los mecanismos más comunes de anticipación es la planificación, en sus variantes de planificación estratégica, planificación de ventas, planificación financiera, planificación de producción, planificación de proyectos, planificación de RRHH y varias otras.

b. Anticipación	Actual	Proyecto
b.1. Modelos de Predicción de Siniestros	Si (solo por monto)	Realizar la predicción a nivel de numero de personas siniestradas por cobertura.
b.2. Predicción de tendencias del mercado.	No	Utilizar data histórica para anticipar cambios de mercado.
b.3. Predecir renovación de clientes	No	Cobrar diferenciado a clientes que existan en los registros

**Tabla 2: Anticipación. Fuente: Elaboración propia.**

#### 4.2.3. Coordinación

La coordinación está estrechamente relacionada con la teoría del mismo nombre, incluyendo otras variables complementarias, además de planificación, tales como el uso de reglas, jerarquía, colaboración y partición. Estas variables deben manejarse de acuerdo a un análisis de costo beneficio. Así, las reglas y la jerarquía son una alternativa más económica a la planificación del punto anterior, permitiendo una coordinación que acepta más recursos de holgura.

c. Coordinación	Actual	Proyecto
c.1. Reglas	Reglas Informales	Reglas formales con apoyo computacional
c.2. Jerarquía	Uso frecuente	Usar solo en caso excepcionales
c.3. Colaboración	Se realiza de manera informal, a través de reuniones.	Formal y con herramientas y sistemas que permiten compartir información relevante
c.4. Partición	No	No

**Tabla 3: Coordinación. Fuente: Elaboración propia.**

#### 4.2.4. Prácticas de Trabajo

Las prácticas de trabajo materializan y detallan las opciones de diseño expresadas en los puntos anteriores. Ellas deben permitir ejecutar las tareas del proceso de manera que se cumpla con tales diseños.

d. Prácticas de Trabajo	Actual	Proyecto
d.1. Lógica de Negocio Automatizada o Semi-automatizada	No	Lógica de pricing Semiautomatizada utilizando DataMining.
d.2. Lógica de Apoyo a Actividades Tácticas	No	Evaluación de Modelos de Pronóstico en base a KPI's
d.3. Procedimientos de Comunicación e Integración	No	Definición de flujo de trabajo para tarificación y calculo de siniestros futuros
d.4. Lógica y Procedimientos de Desempeño y Control	No	Se definen KPI's necesarios para medir desempeño del Proyecto

**Tabla 4: Prácticas de Trabajo. Fuente: Elaboración propia.**

#### 4.2.5. Integración de Procesos Conexos

La integración define el grado de interacción entre los procesos dentro de un Macroproceso o entre diferentes Macroprocesos. Puede tener diversos grados de relación.

e. Integración de Procesos Conexos	Actual	Proyecto
e.1. Proceso Aislado	No	No
e.2. Todos o la Mayor Parte de los Procesos de un Macro Proceso	No	Macroprocesos Administrar relación con el cliente» y «Marketing y análisis de mercado"
e.3. Dos o más Macros que Interactúan	No	No

**Tabla 5: Integración de Procesos Conexos. Fuente: Elaboración propia.**

#### 4.2.6. Mantención Consolidada de Estado

La mantención de estado existe para proveer todos los datos necesarios para ejecutar las prácticas de trabajo y comunicar las actividades y procesos. Esto implica que ellos deben ser ingresados desde los procesos del mismo macro, otros macros que participan, o de sistemas preexistentes, ya sea de la empresa o de otras empresas.

f. Mantención Consolidada de Estado	Actual	Proyecto
f.1. Datos Propios	No	Se creará un DW con datos históricos de SOAP
f.3. Integración con Datos de otros Sistemas de la Empresa	No	Se integrara con datos de VisualTime (Software de Seguros)
f.4. Integración con Datos de Sistemas de otras Empresas	No	No

**Tabla 6: Mantención Consolidada de Estado. Fuente: Elaboración propia.**

#### 4.2.7. Objetivos Específicos

Ya definidos los procesos a rediseñar y las variables correspondientes, se pueden definir los objetivos específicos que serán completados en el desarrollo del proyecto.

- Consolidar información histórica de SOAP en base de datos relacional.
- Mejorar la estimación del OYNR.
- Implementación de algoritmos de segmentación y clasificación a la cartera de clientes del seguro.
- Crear sistemas TI que permitan que el proceso se realice continuamente y en el menor tiempo posible.
- Generar informes de control y gestión del producto de forma periódica.

## Capítulo 5

### Diseño de Procesos

#### 5.1. Diseño BPMN

Dado el análisis de las variables claves y siguiendo la metodología de ingeniería de negocios, se procede a detallar el proceso de “Análisis del comportamiento de clientes” a nivel de BPMN, vale destacar, que el proceso descrito a continuación corresponde a un diseño nuevo, puesto que en la situación actual no existe prácticamente ninguno de los procesos de negocios descritos a continuación.

La notación propia del BPMN puede ser consultada con más detalle en el Anexo B.

##### 5.1.1. Análisis Clientes/Carga DW

El siguiente proceso se encarga principalmente de cargar los nuevos datos (datos mensuales) al Data Warehouse. El proceso incluye una validación física y lógica para los datos brutos de operaciones y luego se realiza un ETL, para poder utilizar los datos de mejor forma en los procesos siguientes.

Los datos provienen de distintas fuentes y necesitan ser limpiados y estandarizados, para esto se utiliza el ETL, que es construido en Pentaho y permite la automatización de parte importante del proceso.

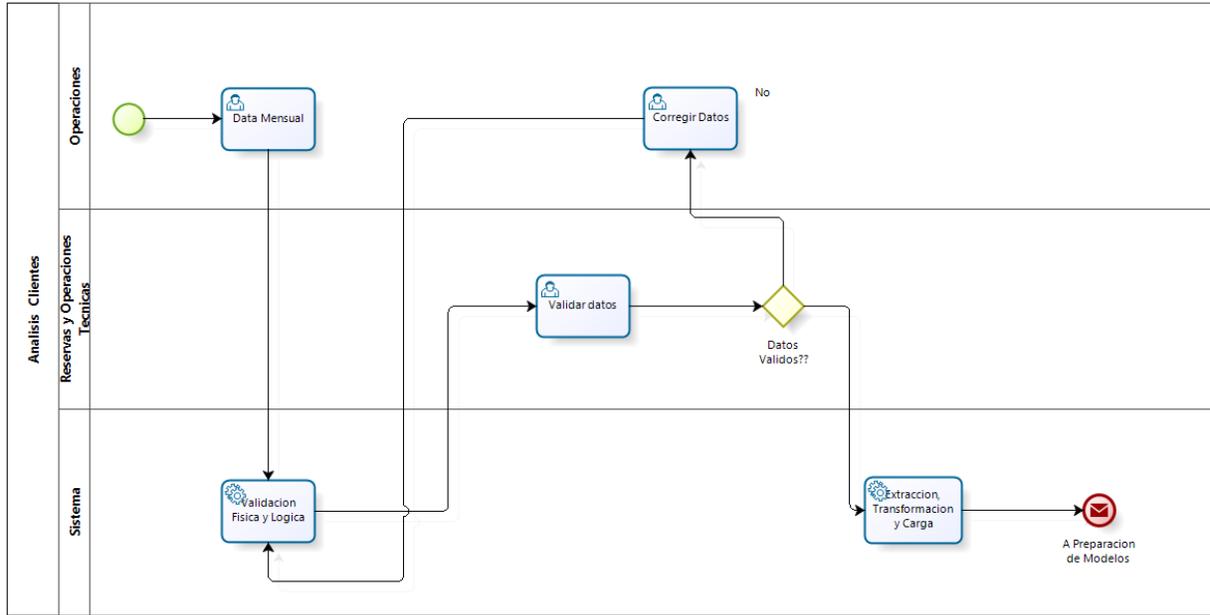


Figura 23: Carga Date Warehouse. Fuente: Elaboración propia.

### 5.1.1.1. Automatización

Con respecto a la automatización, se optó por utilizar el software Pentaho, debido a sus capacidades como ETL (antes llamado Kettle) y que además es un de código libre.

El proceso en general radica en la adquisición de la data necesaria proveniente de distintas fuentes, su posterior transformación y finalmente la consolidación de los datos en el respectivo Datawarehouse. La información se puede dividir principalmente en dos categorías; producción y siniestros.

La producción corresponde a todos los datos que caracterizan a las pólizas (características tanto del dueño del seguro como del vehículo asegurado), mientras que los siniestros simplemente reflejan el monto de los siniestros (montos pagados por la empresa a los asegurados) por cobertura. Estos archivos son generados tanto en Access, como en Excel, lo que además se complementa con algunos archivos TXT (información del Servel). Toda esta información termina en bases de datos PostgreSQL (Data Warehouse).

A continuación se aprecia el proceso general construido en Pentaho. Algunos de los subprocessos claves serán detallados más adelante. Como resumen, los datos de producción se validan y de no haber error se cargan el DW, para luego actualizar varios campos mediante instrucciones SQL. Esto mismo pasa con los siniestros, aunque aquí se lee la tabla resultante del proceso de cálculo de OYNR.

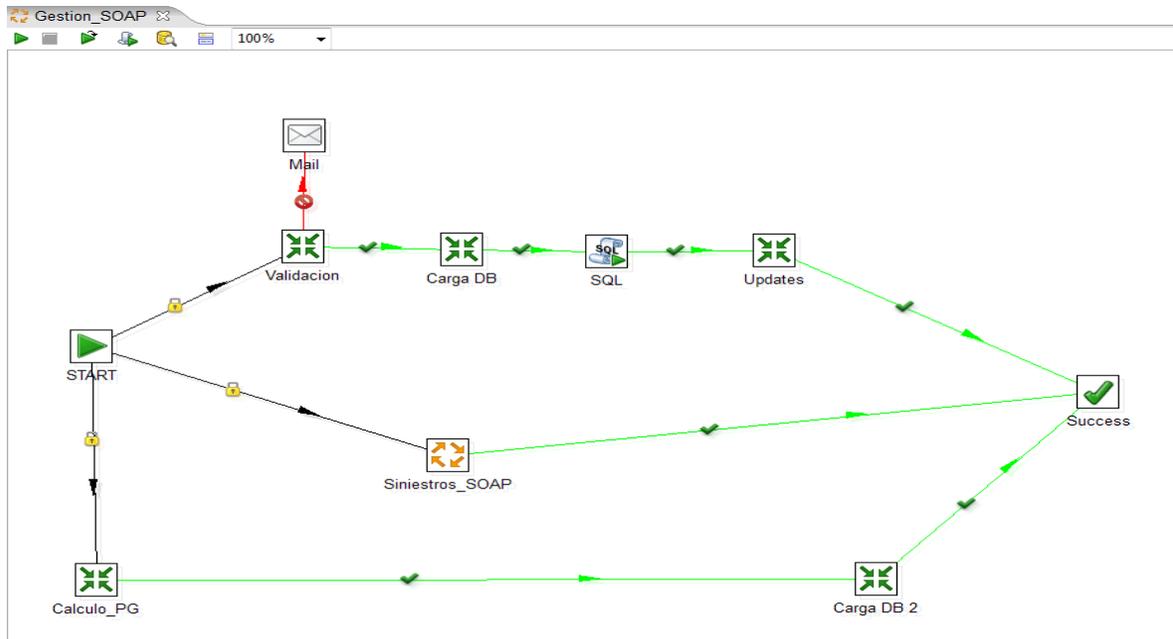


Figura 24: Proceso general en Pentaho. Fuente: Elaboración Propia.

La figura siguiente muestra el proceso de validación, este es sumamente expedito (debido a la forma en que Pentaho procesa las filas de datos) y genera (en el caso de ser necesario) un reporte de errores (en Excel) que se envía inmediatamente al responsable de los datos para su pronta corrección. Vale destacar, que si alguna de las validaciones no se cumplen, el proceso total es abortado, por lo que nada pasa esta etapa al menos que todo este validado. En el caso de estar todo validado el proceso sigue de forma autónoma, cargando y actualizando el Datawarehouse para futuros análisis.

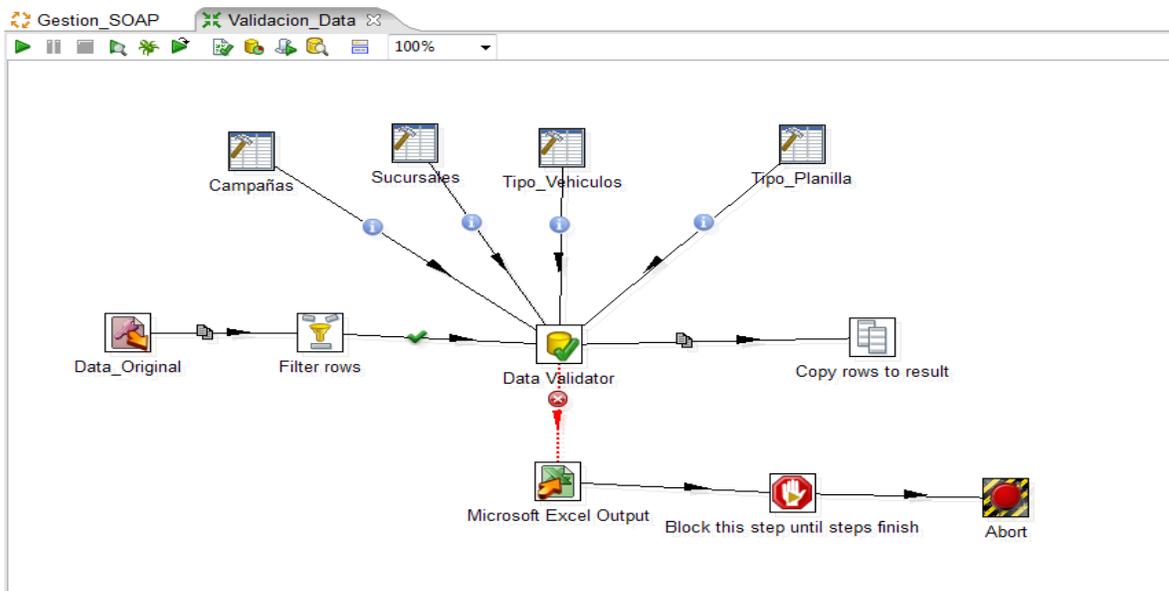


Figura 25: Validación de datos en Pentaho. Fuente: Elaboración Propia.

Luego de la validación comienza el proceso de carga antes descrito (figuras 35 y 36). El resultante del proceso de validación es cargado automáticamente al Datawarehouse en PostgreSQL para luego actualizar los campos necesarios con respecto a otra información relevante (Updates). Esto es por ejemplo, asignar los canales de venta, la comisión correspondiente a cada póliza y finalmente la actualización de la región y comuna mediante la información obtenida de parte del Serval (además de poder obtener el sexo de cada asegurado).

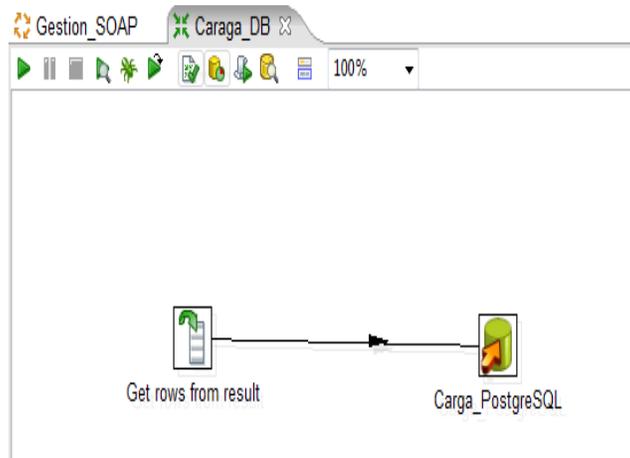


Figura 26: Carga DW en Pentaho. Fuente: Elaboración Propia.

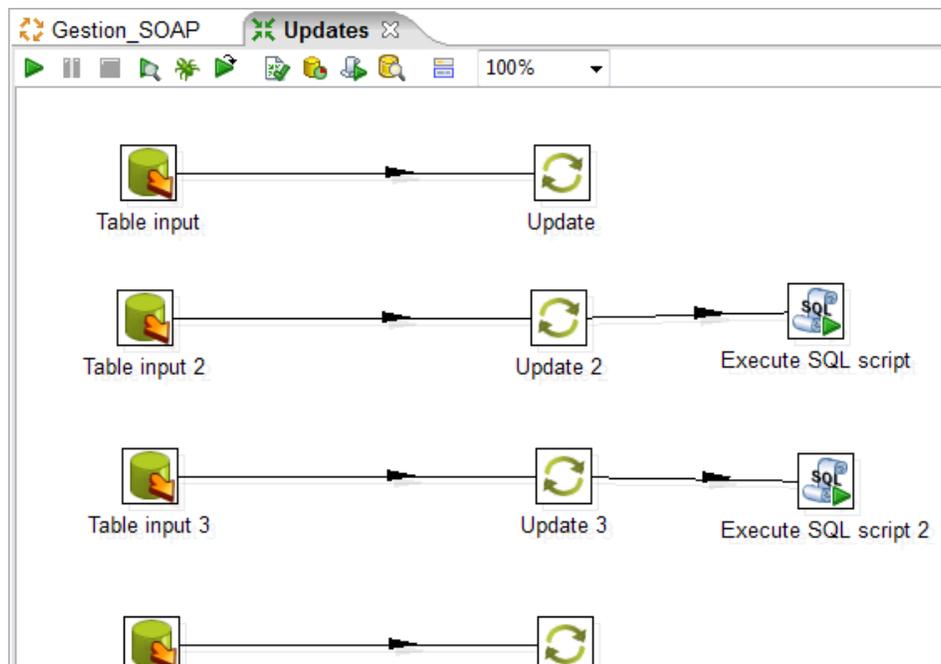
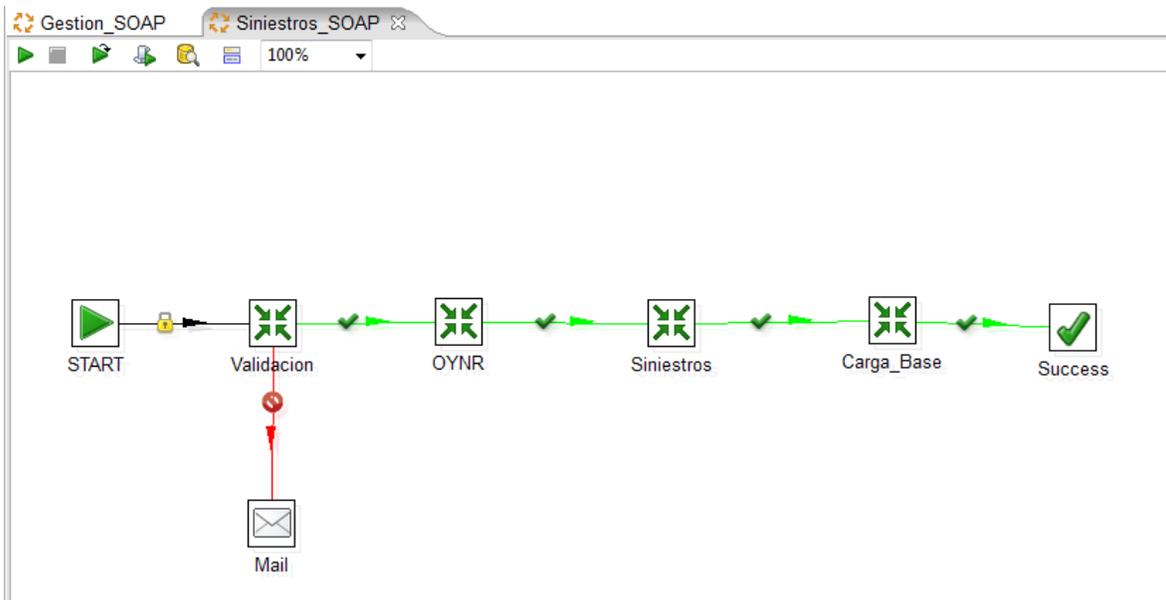


Figura 27: Updates en Pentaho. Fuente: Elaboración Propia.

El tratamiento de los siniestros es un proceso aparte (como se puede de distinto color en el proceso general). El proceso inicia con validación de los mismos, la validación es menos exhaustiva que la del proceso de producción ya que en realidad se utilizan solo 2 datos, el monto del siniestro final y a que cobertura se asocia el pago del mismo.

Con los siniestros validados, se entrega el archivo al proceso de OYNR que calcula este valor y lo actualiza en la tabla de siniestros. Luego se crean tablas auxiliares agrupadas por número de siniestros y número de afectados por siniestro (este proceso se detalla posteriormente).

Finalmente toda esta información es cargada al Datawarehouse, con lo que se tiene todas las características de las pólizas, así como las primas y los siniestros correspondientes. Con esta información ya se puede calcular el resultado financiero del seguro y además ya se pueden aplicar los algoritmos de Data Mining previamente escogidos.



**Figura 28: Procesamiento Siniestros en Pentaho. Fuente: Elaboración Propia.**

Durante la carga del data Warehouse histórico se utilizó mucho tiempo intentando normalizar los datos y eliminar la mayor cantidad de errores o datos faltantes. Muchos errores provienen de que muchas pólizas de SOAP son vendidas a “mano” por lo que los agentes rellenan la información solicitada con lápiz y papel, lo que produce muchos errores involuntarios.

Con ayuda de los departamentos de operaciones de seguros generales se logró corregir muchos datos que presentaban falencias (estas correcciones se centraron en las variables que fueron utilizadas

finalmente en el análisis). Para los datos perdidos (MCAR<sup>6</sup>, MAR<sup>7</sup>, MNAR<sup>8</sup>) se ocuparon varias técnicas comunes en el procesamiento de datos tal como remplazo por medias, modas, utilización de regresiones, etc.

Este proceso se intentó automatizar para la carga mensual de datos, para que no se necesitara de intervención humana tan constante. Un caso de especial importancia en la carga de los nuevos datos fue el de la variable “Marca”, esto debido a que a priori el equipo supuso que podría tener un impacto interesante en los modelos (además de no estar siendo utilizada en el momento del desarrollo del presente proyecto).

Como la variable “Marca” fue una de la que presento mayor número de errores (principalmente porque muchas personas no saben cómo escribir el nombre de ciertas marcas extranjeras de forma correcta), se tuvo que idear un método para poder cargar los datos con el menor número de errores posibles y de forma automática.

El primer paso fue conseguir con el departamento de vehículos una lista con todas las marcas comercializadas en el país (bien escritas y actualizado de forma periódica). Luego se creó un proceso en donde cada nuevo registro busca si la marca que viene es igual a una de las marcas conocidas, de ser así, el registro entra en el proceso normal de carga descrito anteriormente. En caso de no tener una correspondencia directa, se calcula el indicador GED (General Edit Distance) [27] con todas las marcas conocidas. La técnica conocida como GED se utiliza para compararse dos cadenas y calcula el costo mínimo de una secuencia de operaciones para construir la cadena2 a partir de la cadena1. En este cálculo, distintas operaciones consideran distintas penalizaciones, por ejemplo insertar o eliminar un carácter para pasar de la cadena1 a la cadena2 es penalizado con 100 puntos, pero diferencias de solo un espacio en blanco conlleva solo 10. Utilizando esta técnica se puede ranquear de buena forma cual es la marca más probable a la que hace referencia el texto ingresado (la que tenga la menor distancia). Luego simplemente el texto original es remplazado por la marca más probable y el proceso continúa de manera normal. Por razones obvias este método no funciona cuando no viene la información de la marca, pero estos son casos muy raros y son resueltos de otras formas (en general se solicita la corrección del dato al área correspondiente).

### 5.1.2. Análisis Clientes/Preparación de Modelos

Luego de la carga del Datawarehouse, se procede a la preparación de modelos, en donde se realizan algunas validaciones más profundas (que no pueden automatizarse dentro del ETL), para luego realizar el cálculo de OYNR y dejar todos los datos preparados para la aplicación de cualquier algoritmo que se escoja.

Cabe destacar, que en el proceso de “Consolidación de información”, además de la consolidación de los datos con los provenientes del cálculo de OYNR se entrena el algoritmo de la clasificación que luego será ejecutado en el canal web para predecir el comportamiento de los clientes y otorgar descuentos acordes.

---

<sup>6</sup> MCAR = Missing completely at random

<sup>7</sup> MAR = Missing at random

<sup>8</sup> MNAR = Missing not at random

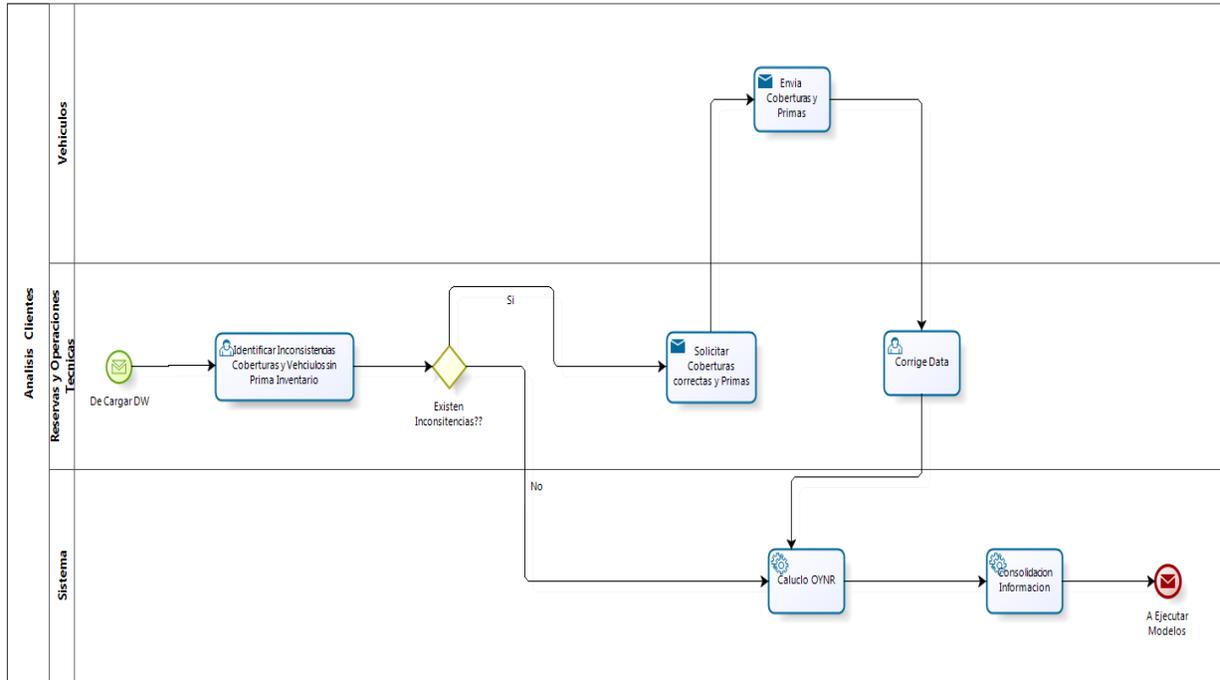


Figura 29: Preparación de Modelos. Fuente: Elaboración propia.

5.1.2.1. OYNR

El termino OYNR corresponde a los siniestros “Ocurridos Y No Reportados” es decir, siniestros que ya ocurrieron pero todavía no han sido reclamados, este es un valor crítico, ya que se debe generar reserva extra por el número de OYNR y por ende al subestimar los OYNR se está generando una utilidad falsa y en caso de sobreestimarlos se está generando una perdida contable (que no es real pero aparecerá en los estados financieros).

La superintendencia de valores y seguros (SVS) norma que las empresas de seguros utilicen el método convencional llamado “Chain Ladder” para la estimación de OYNR (ejemplo en figura siguiente), un método que usa datos históricos para estimar factores que luego multiplican los siniestros pagados y de esta forma estiman los siniestros que ya han ocurrido pero llegarán con cierto desfase.

Accident year <i>i</i>	Developing year <i>j</i>									
	1	2	3	4	5	6	7	8	9	10
1981	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834
1982	106	4285	5396	10666	13782	15599	15496	16169	16704	
1983	3410	8992	13873	16141	18735	22214	22863	23466		
1984	5655	11555	15766	21266	23425	26083	27067			
1985	1092	9565	15836	22169	25955	26180				
1986	1513	6445	11702	12935	15852					
1987	557	4020	10946	12314						
1988	1351	6947	13112							
1989	3133	5395								
1990	2063									

Figura 30: Ejemplo Chain Ladder [11].

Este método tiene dos supuestos fuertes que no siempre se cumplen, estos son; que los pagos mes a mes son independientes y que los desfases son constantes en el tiempo. Muchas veces esto no se cumple lo que provoca grandes diferencias entre la estimación y la realidad.

Además del problema en la estimación, en la actualidad el proceso no está automatizado, lo que conlleva una cantidad de tiempo considerable en su cálculo.

Teniendo en cuenta lo anterior, el proyecto contempla que el sistema TI propuesto calcule el OYNR, utilizando un método que obtenga resultados considerablemente mejores que los del método tradicional.

Para esto se utilizó una modificación del método “Chain Ladder”, en donde se parte por utilizar la forma mensual y luego se abre por cobertura, es decir, se analizan los siniestros (y se construyen los triángulos) con respecto a las coberturas del seguro (gastos médicos, invalidez y muerte). Luego se calculan los factores pero los que presentan un OYNR mayor al 20% de los siniestros incurridos son ajustados utilizando la frecuencia y el costo medio de los últimos 12 meses móviles.

En las siguientes figuras se aprecia parte de la construcción del modelo “Chain Ladder modificado”.

INCIDENTOS		Dólar																																																																																																				
Evoluciones		1	1	1	2	2	2	3	3	3	4	4	4	5	5	5	6	6	6	7	7	7	8	8	8	9	9	9	10	10	10	11	11	11	12	12	12	13	13	13	14	14	14	15	15	15																																																								
1	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

Figura 31: Calculo de OYNR (1). Fuente: Elaboración Propia.

Agrupación	Mes	Fecha	Proceso	Pagado	No ret	No sin	UP No	UP UP	UP No	UP UP	OYNR	CM	Frecuen	Flag	Flag	Metodo	Dias	Espeor	N sin	N sin	Costo	Costo	Costo	Costo	
			date	UP	ocurridos	ocurridos	Sin	Sin	Sin	Sin		ajustado	O Ajustada	CM	Frecuen	Log	Ganados	costo	año a	año a	medio	medio	medio	medio	
																			cada mil	cada mil	(movil 12	(movil 12	(movil 12	(movil 12	
GRMUTUP	2	01-02-2009	200902	4.732	95,00	95,00	192	4.732	192	4.732	-	31	297	-	-	-	5.696.626	196.688	297	297	311	297	311	297	311
GRMUTUP	3	01-03-2009	200903	4.995	95,00	95,00	193	4.995	193	4.995	-	33	267	-	-	-	6.275.689	209.190	267	267	326	267	326	267	326
GRMUTUP	4	01-04-2009	200904	5.097	146,00	146,00	145	5.097	145	5.097	-	27	276	-	-	-	6.905.841	193.520	276	276	311	276	311	276	311
GRMUTUP	5	01-05-2009	200905	6.050	144,00	144,00	144	6.050	144	6.050	-	42	299	-	-	-	6.090.191	203.006	299	299	420	299	420	299	420
GRMUTUP	6	01-06-2009	200906	5.246	130,00	130,00	130	5.246	130	5.246	-	40	242	-	-	-	5.690.367	196.240	242	242	404	242	404	242	404
GRMUTUP	7	01-07-2009	200907	4.976	139,00	139,00	139	4.976	139	4.976	-	25	249	-	-	-	6.143.956	204.244	249	249	261	249	261	249	261
GRMUTUP	8	01-08-2009	200908	5.056	141,00	141,00	141	5.056	141	5.056	-	36	248	-	-	-	6.225.894	207.863	248	248	359	248	359	248	359
GRMUTUP	9	01-09-2009	200909	4.445	146,00	146,00	145	4.445	145	4.445	-	32	226	-	-	-	6.193.964	203.862	226	226	233	226	233	226	233
GRMUTUP	10	01-10-2009	200910	2.422	136,00	136,00	136	2.422	136	2.422	-	19	217	-	-	-	6.364.559	242.862	217	217	192	217	192	217	192
GRMUTUP	11	01-11-2009	200911	2.899	125,00	125,00	125	2.899	125	2.899	-	23	221	-	-	-	6.391.044	206.956	221	221	233	221	233	221	233
GRMUTUP	12	01-12-2009	200912	3.095	124,00	124,00	124	3.095	124	3.095	-	24	232	-	-	-	6.499.003	213.653	232	232	243	232	243	232	243
GRMUTUP	1	01-01-2010	201001	3.170	139,00	139,00	139	3.170	139	3.170	-	23	227	-	-	-	6.430.766	214.259	227	227	245	227	245	227	245
GRMUTUP	2	01-02-2010	201002	3.422	139,00	139,00	139	3.422	139	3.422	-	21	223	-	-	-	6.519.446	194.676	223	223	239	223	239	223	239
GRMUTUP	3	01-03-2010	201003	4.987	139,00	139,00	139	4.987	139	4.987	-	36	234	-	-	-	6.509.958	226.867	234	234	358	234	358	234	358
GRMUTUP	4	01-04-2010	201004	2.253	132,00	132,00	132	2.253	132	2.253	-	21	222	-	-	-	6.509.958	226.867	222	222	239	222	239	222	239
GRMUTUP	5	01-05-2010	201005	2.598	97,00	97,00	97	2.598	97	2.598	-	27	213	-	-	-	6.509.958	226.867	213	213	239	213	239	213	239
GRMUTUP	6	01-06-2010	201006	2.550	110,00	110,00	110	2.550	110	2.550	-	23	254	-	-	-	6.509.958	226.867	254	254	239	254	239	254	239
GRMUTUP	7	01-07-2010	201007	3.206	98,00	98,00	98	3.206	98	3.206	-	32	211	-	-	-	6.509.958	226.867	211	211	239	211	239	211	239
GRMUTUP	8	01-08-2010	201008	3.169	129,00	129,00	129	3.169	129	3.169	-	25	275	-	-	-	6.509.958	226.867	275	275	239	275	239	275	239
GRMUTUP	9	01-09-2010	201009	2.211	105,00	105,00	105	2.211	105	2.211	-	21	229	-	-	-	6.509.958	226.867	229	229	239	229	239	229	239
GRMUTUP	10	01-10-2010	201010	1.886	100,00	100,00	100	1.886	100	1.886	-	17	230	-	-	-	6.509.958	226.867	230	230	239	230	239	230	239
GRMUTUP	11	01-11-2010	201011	2.489	117,00	117,00	117	2.489	117	2.489	-	21	237	-	-	-	6.509.958	226.867	237	237	239	237	239	237	239
GRMUTUP	12	01-12-2010	201012	1.232	98,00	98,00	98	1.232	98	1.232	-	12	190	-	-	-	6.509.958	226.867	190	190	239	190	239	190	239
GRMUTUP	1	01-01-2011	201101	2.397	131,00	131,00	131	2.397	131	2.397	-	17	251	-	-	-	6.509.958	226.867	251	251	239	251	239	251	239
GRMUTUP	2	01-02-2011	201102	2.797	146,00	146,00	146	2.797	146	2.797	-	26	205	-	-	-	6.509.958	226.867	205	205	239	205	239	205	239
GRMUTUP	3	01-03-2011	201103	1.491	101,00	101,00	101	1.491	101	1.491	-	15	189	-	-	-	6.509.958	226.867	189	189	239	189	239	189	239
GRMUTUP	4	01-04-2011	201104	2.033	114,00	114,00	114	2.033	114	2.033	-	18	194	-	-	-	6.509.958	226.867	194	194	239	194	239	194	239
GRMUTUP	5	01-05-2011	201105	2.494	95,00	95,00	95	2.494	95	2.494	-	28	194	-	-	-	6.509.958	226.867	194	194	239	194	239	194	239
GRMUTUP	6	01-06-2011	201106	4.297	136,00	136,00	136	4.297	136	4.297	-	32	224	-	-	-	6.509.958	226.867	224	224	239	224	239	224	239
GRMUTUP	7	01-07-2011	201107	1.662	129,00	129,00	129	1.662	129	1.662	-	13	204	-	-	-	6.509.958	226.867	204	204	239	204	239	204	239
GRMUTUP	8	01-08-2011	201108	2.186	124,00	124,00	124	2.186	124	2.186	-	18	194	-	-	-	6.509.958	226.867	194	194	239	194	239	194	239
GRMUTUP	9	01-09-2011	201109	2.391	137,00	137,00	137	2.391	137	2.391	-	16	220	-	-	-	6.509.958	226.867	220	220	239	220	239	220	239
GRMUTUP	10	01-10-2011	201110	1.896	110,00	110,00	110	1.896	110	1.896	-	17	175	-	-	-	6.509.958	226.867	175	175	239	175	239	175	239
GRMUTUP	11	01-11-2011	201111	2.430	132,00	132,00	132	2.430	132	2.430	-	7	183	-	-	-	6.509.958	226.867	183	183	239	183	239	183	239
GRMUTUP	12	01-12-2011	201112	1.929	130,00	130,00	130	1.929	130	1.929	-	12	209	-	-	-	6.509.958	226.867	209	209	239	209	239	209	239
GRMUTUP	1	01-01-2012	201201	2.002	136,00	136,00	136	2.002	136	2.002	-	12	195	-	-	-	6.509.958	226.867	195	195	239	195	239	195	239
GRMUTUP	2	01-02-2012	201202	3.279	161,00	161,00	161	3.279	161	3.279	-	22	201	-	-	-	6.509.958	226.867	201	201	239	201	239	201	239
GRMUTUP	3	01-03-2012	201203	2.020	145,00	145,00	145	2.020	145	2.020	-	21	217	-	-	-	6.509.958	226.867	217	217	239	217	239	217	239
GRMUTUP	4	01-04-2012	201204	4.899	246,00	246,00	246	4.899	246	4.899	-	33	188	-	-	-	6.509.958	226.867	188	188	239	188	239	188	239
GRMUTUP	5	01-05-2012	201205	2.747	230,00	230,00	230	2.747	230	2.747	-	29	207	-	-	-	6.509.958	226.867	207	207	239	207	239	207	239
GRMUTUP	6	01-06-2012	201206	4.097	235,00	235,00	235	4.097	235	4.097	-	34	177	-	-	-	6.509.958	226.867	177	177	239	177	239	177	239
GRMUTUP	7	01-07-2012	201207	7.053	230,00	230,00	230	7.053	230	7.053	-	24	233	-	-	-	6.509.958	226.867	233	233	239	233	239	233	239
GRMUTUP	8	01-08-2012	201208	4.581	262,00	262,00	262	4.581	262	4.581	-	55	179	-	-	-	6.509.958	226.867	179	179	239	179	239	179	239
GRMUTUP	9	01-09-2012	201209	4.000	261,00	261,00	262	4.000	262	4.000	-	52	193	-	-	-	6.509.958	226.867	193	193	239	193	239	193	239
GRMUTUP	10	01-10-2012	201210	5.267	261,00	261,00	262	5.267	262	5.267	-	46	177	-	-	-	6.509.958	226.867	177	177	239	177	239	177	239
GRMUTUP	11	01-11-2012	201211	4.262																					

### 5.1.3. Análisis Clientes/Ejecución de Modelos

El proceso final del rediseño corresponde a la ejecución de la segmentación y el análisis de las variables claves correspondiente (además de poder obtener la prima de riesgo asociada a cada segmento).

Con todos los datos preparados desde el proceso “Preparación de modelos”, se procede a segmentar utilizando algún algoritmo de Clustering apropiado.

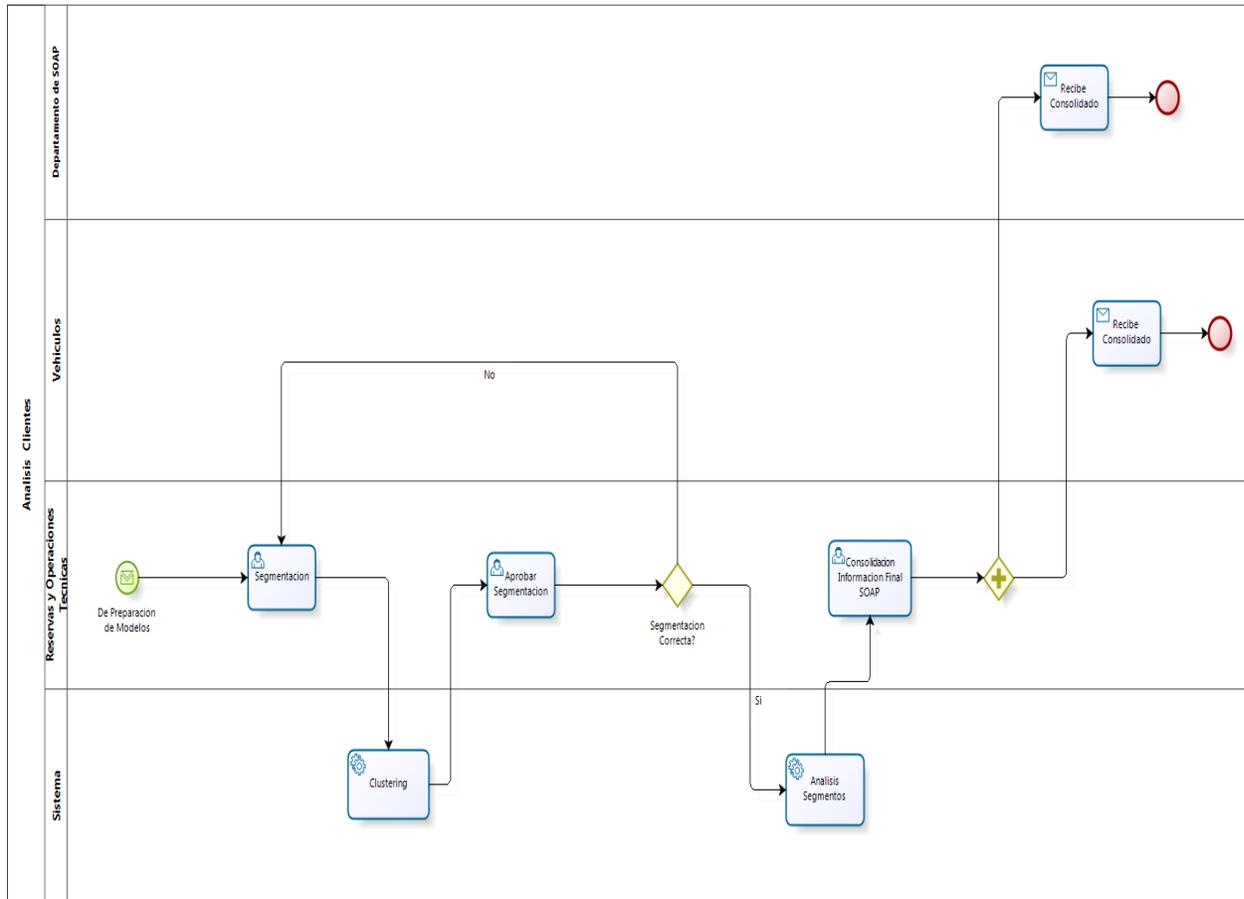


Figura 33: Ejecución de Modelos. Fuente: Elaboración propia.

#### 5.1.3.1. Clustering

Dada la literatura [20] para el proceso de Clustering se escogieron dos algoritmos para analizar; el algoritmo K-means y conglomerado bietapico, principalmente por sus buenos resultados con grandes cantidades de datos y la posibilidad de manejar variables tanto numéricas como categóricas.

### 5.1.3.2. K-Means

K-means es un método de agrupamiento, que tiene como objetivo la partición de un conjunto de  $n$  observaciones en  $k$  grupos en el que cada observación pertenece al grupo más cercano a la media [22].

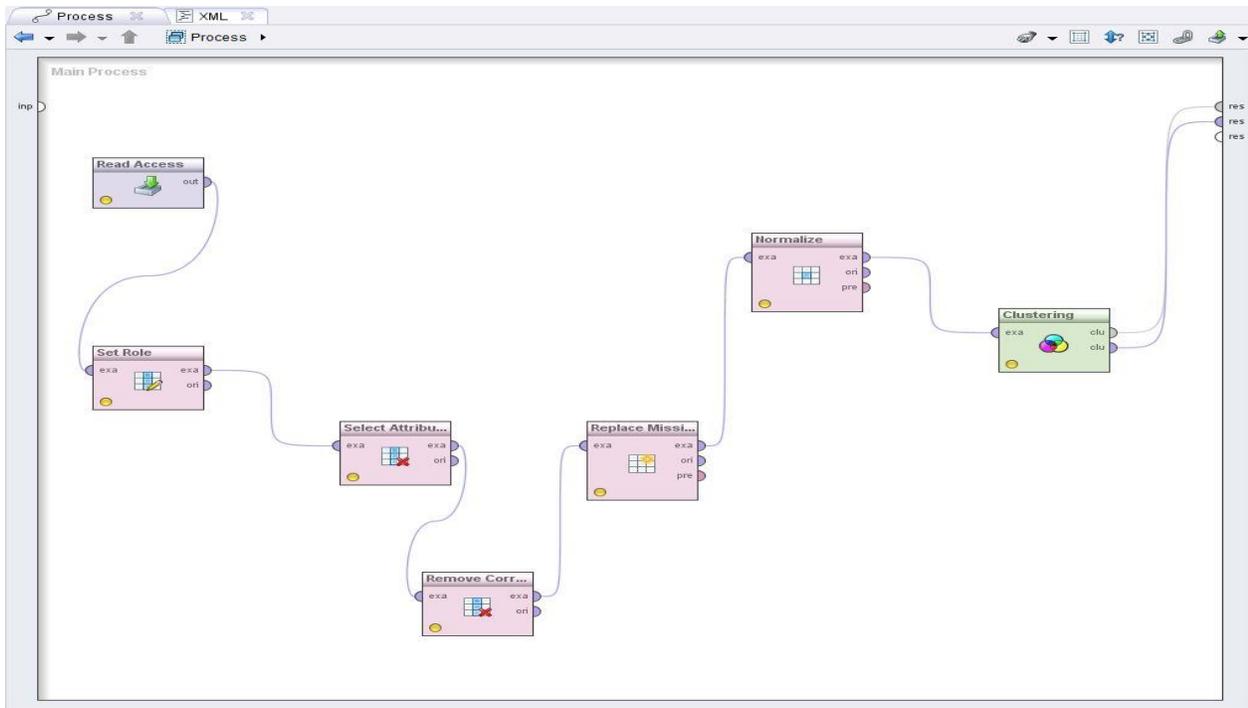
Para determinar el número de clusters se calcularon 3 indicadores para distintos número de clusters; el indicador de Ginni, Davies-Bouldin y suma de cuadrados. En relación a los clúster cada indicador representa una mejoría al tomar un valor menor, por esta razón se decide escoger 5 clúster. Esta decisión se fundamenta principalmente en el indicador Davies-Bouldin que al ser menor representa una menor distancia intra-cluster y una mayor distancia extra-clúster.

Indicador/Nº Clusters	3	5	7	9	11
Ginni	0,99975	0,99874	0,99675	0,99628	0,99503
Davies-Bouldin	1,068	<b>0,943</b>	0,984	1,050	1,029
Sum Of Squares	0,448	0,310	0,193	0,151	0,137

**Tabla 8: Indicadores Número de Clusters. Fuente: Elaboración Propia.**

En la siguiente figura se puede visualizar el proceso creado en Rapidminer que se encarga de realizar el Clustering a la base de datos histórica ya consolidada de SOAP. Vale destacar que aunque originalmente K-means no puede manejar variables categóricas, Rapidminer tiene una versión que si lo maneja, al utilizar distintas medidas de distancias tanto para variables numéricas como nominales.

A pesar de tener que manejar un par de millones de datos, el modelo utilizado, tiene una buena performance ya que se demora alrededor de 10 a 12 minutos en realizar el Clustering (corriendo en una maquina con solo 3 GB de ram).



**Figura 34: Modelo Clustering en RapidMiner. Fuente: Elaboración Propia.**

Las siguientes tablas muestran algunos resultados del Clustering realizado con el algoritmo K-Means.

Conglomerado	1	2	3	4	5
1		28,189	24,532	24,279	25,398
2	28,189		20,735	16,107	14,888
3	24,532	20,735		16,933	12,613
4	24,279	16,107	16,933		21,663
5	25,398	14,888	12,613	21,663	

**Tabla 9: Distancias entre los centros de los conglomerados finales. Fuente: Elaboración Propia.**

En la tabla anterior, se aprecian las distancias entre los distintos clusters, lo que permite apreciar, por ejemplo, que el Clúster 1 y el 2 son los que están más alejados, mientras que el 2 con el 5 son los más cercanos, por lo que sus comportamientos deberían estar acorde a estas distancias.

La siguiente tabla, muestra los estadísticos típicos del análisis ANOVA, lo que nos entrega información muy valiosa sobre la cartera analizada. Primeramente se observa que todas las variables son utilizadas para realizar los clusters, además el estadístico “F” nos dice cuan relevante fueron en la formación de los mismos. El tipo de vehículo resulta ser la variable más discriminante, lo que al parecer no es muy sorprendente, ya que la diferencia entre un Jeep y una moto es bastante grande. Lo que resulta

bastante interesante es que la región o la comuna son variables mucho menos relevantes que la Marca y la antigüedad del vehículo. Estas variables en este momento no son utilizadas en la tarificación y que serán entregadas por primera vez como producto del presente proyecto.

	Conglomerado		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
COD_REGION	11739,216	4	12,630	246489	929,438	,000
COD_SUBGRUPO_REGION	981,302	4	,790	246489	1241,799	,000
COD_TIPO_VEHICULO_TECNICO	5632207,822	4	14,226	246489	395922,582	,000
Cod_Marca	2259834,150	4	23,098	246489	97836,264	,000
ANTIGUEDAD_Auto	1986016,356	4	21,473	246489	92490,303	,000
Cod_Sexo	801,011	4	,462	246489	1735,089	,000

Tabla 10: ANOVA Conglomerados K-Means. Fuente: Elaboración Propia.

### 5.1.3.3. Conglomerado Bietapico

Este algoritmo se recomienda cuando las muestras son muy grandes (como es el caso) y además tiene la posibilidad de dar tratamiento separado a las variables categóricas y a las continuas.

Para poder comparar de mejor forma con el algoritmo anterior, también se escogieron 5 segmentos, por un lado el tiempo de respuesta fue mayor que el de K-means, aunque no de manera decisiva. Se calculó el coeficiente Silhouette [28], el cual representa la cohesión y separación de un cluster. Un coeficiente cercano a +1 implica que los puntos son cercanos a su propio cluster y lejanos a los otros clusters.

El coeficiente Silhouette se puede definir de la siguiente forma, para cada punto ( $x_i$ ) el coeficiente  $s_i$  es el siguiente:

$$s_i = \frac{\mu_{out}^{\min}(x_i) - \mu_{in}(x_i)}{\max\{\mu_{out}^{\min}(x_i), \mu_{in}(x_i)\}}$$

Donde,

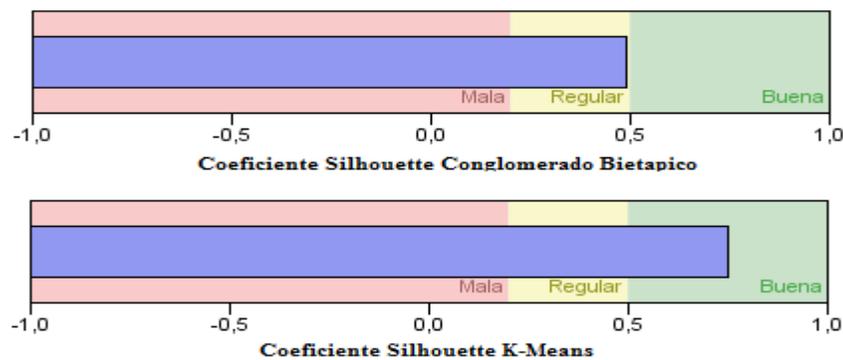
$\mu_{in}(x_i)$  es la distancia promedio de  $x_i$  a los puntos en su mismo cluster.

$\mu_{out}^{\min}(x_i)$  es la distancia promedio de  $x_i$  a los puntos del clusters más cercano.

Dado lo anterior el coeficiente Silhouette (SC) es el promedio de todos los  $s_i$ , es decir:

$$SC = \frac{1}{n} \sum_{i=1}^n s_i$$

Realizando el cálculo descrito es que se pudo determinar el coeficiente Silhouette para ambos algoritmos, los que se presentan en la siguiente figura:



**Figura 35: Comparación Algoritmos Clustering. Fuente: Elaboración Propia.**

Por las razones antes expuestas es que se escoge el algoritmo K-Means y los siguientes análisis se hacen con respecto a sus resultados.

Finalmente se consolida toda la información del proceso, en donde se genera un informe de que es entregado tanto al área de vehículos (responsables de la tarificación) como al área de SOAP (área responsable de la comercialización del producto). Este informe consiste de varias partes, por una parte está la información financiera y contable del producto, en donde es fácil visualizar si se está ganando o perdiendo dinero y cuál es la razón. Además se genera una vista (principalmente para el departamento de vehículos) con los precios teóricos de los segmentos encontrados, esta información la utilizaran para mejorar sus proceso de tarificación para las siguientes campañas. Por otro lado también se presentan las variables más relevantes que expliquen la siniestralidad de la cartera (relación entre la prima y los siniestros) y sub-segmentos que presenten comportamientos interesantes, esta información la utilizara el departamento de SOAP para enfocar mejor sus alianzas y su forma de atacar los targets escogidos (gastos de marketing, elección de intermediario, canales de comercialización, etc.).

Como se verá más adelante, este proceso se puede repetir todas las veces que sea necesaria, según los deseos del ejecutante, esto difiere de los procesos anteriores, ya que, sería extraño que se ejecutaran muchas veces dentro del mismo mes.

#### 5.1.4. Venta y Atención Clientes

En general es imposible que el rediseño de un proceso no modifique en algo los procesos contiguos, y este caso no es una excepción, ya que, al menos debería modificar los procesos de venta y de decisiones de marketing.

Las decisiones de marketing se escapan al alcance del proyecto (puesto que, dependen de muchas variables que no se manejan) mientras que las ventas si se puede rediseñar en relaciones a las nuevas posibilidades otorgadas por la ejecución del nuevo análisis de clientes.

La venta de SOAP se realiza de muchas formas distintas; en las distintas sucursales de la empresa, en stand puestos en lugares estratégicos (como municipalidades), mediante intermediarios (que ellos se encargan de venderlo tanto a personas como a empresas) y en formato web a través de la página de la empresa. Debido a la dificultad propia de implementar los sistemas TI correspondientes en la mayoría de las situaciones descritas con anterioridad, es que se escoge la venta por el Canal web (ver anexo C), ya que se pueden solicitar todos los datos necesarios para el modelo de clasificación y además la comisión es 0 (lo que es muy relevante a la hora de ver el margen final del producto).

El cambio más importante en la venta mediante el canal web, es que ahora se utilizara un algoritmo para clasificar a los clientes según sus características más relevantes (las cuales se desprenden del análisis del Clustering), por lo que se obtendrán probabilidades de que el cliente sea Malo, Riesgoso o Bueno y por ende se le podrá otorgar un descuento a los clientes que tengan una alta probabilidad de ser buenos. No se generara un recargo especial para quienes tengan una mayor probabilidad de caer en los estados restantes, ya que la política de la empresa implica no recargar más allá de la prima de riesgo calculada en la segmentación.

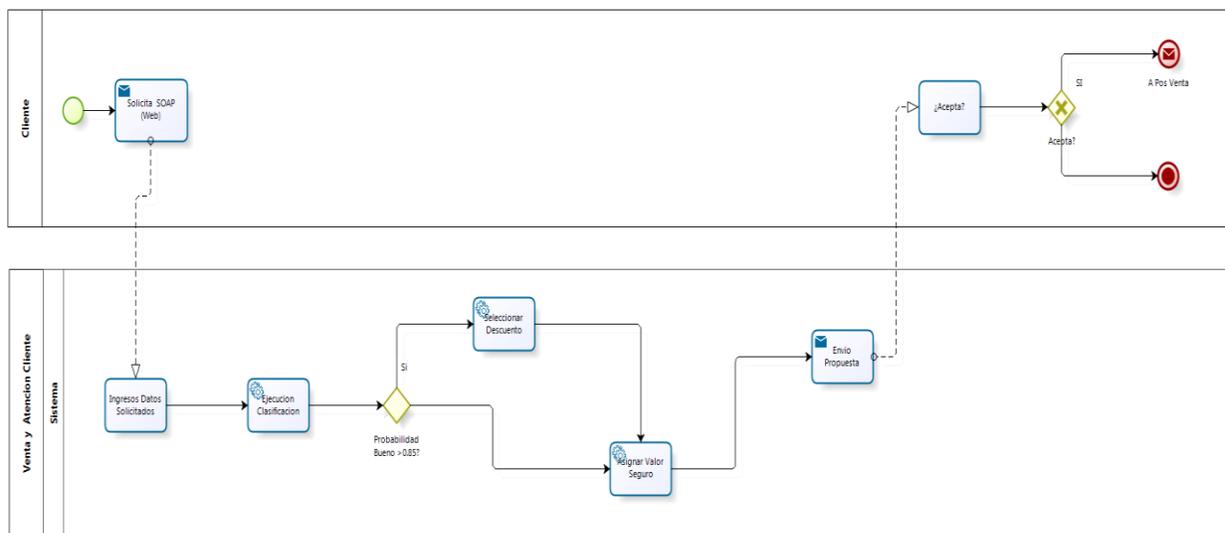


Figura 36: Venta y Atención Clientes. Fuente: Elaboración propia.

Para poder clasificar a los clientes se escoge el siguiente algoritmo, detallado a continuación.

#### 5.1.4.1. Clasificación

El proceso de clasificación de clientes, está asociado a la venta por el canal web, donde dependiendo de las características ingresadas, se clasificara al cliente (por medio de probabilidades), en Malo (M), Riesgoso (R) o Bueno (B), con lo que se le otorgara un descuento en su prima.

Para crear estas categorías se utilizó la información obtenida en el proceso de Clustering, en donde se detectó que las principales variables que explican el comportamiento de los clientes en SOAP son: “Tipo de Vehículo”, “Marca Vehículo” y “Antigüedad Vehículo”. Sabiendo esto se analizó estas variables con respecto a su costo medio y frecuencia de siniestros para escoger que valores poblarían que categoría (M, R o B). De esta forma se obtuvo la siguiente distribución porcentual.

Categorías	Nº Pólizas
Malo	0,4%
Riesgoso	75,8%
Bueno	23,8%

**Tabla 11: Distribución Categorías Clasificación. Fuente: Elaboración propia**

Vale destacar, que luego de la clasificación solo se le aplicará el descuento a las pólizas que tengan una probabilidad mayor al 80% de ser clasificadas como buenas.

Debido a la amplia gama de algoritmos capaces de realizar la clasificación necesaria, se optó por preseleccionar solo dos de ellos debidos a que según la literatura eran más compatibles con el problema tratado (cantidad de datos, dimensionalidad, etc.) [23]. Estos algoritmos son: “Redes Neuronales (Perceptron multicapa) y Arboles de Decisión”.

#### 5.1.4.2. Redes Neuronales

Las redes de neuronas artificiales (denominadas habitualmente como RNA) son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los animales. Se trata de un sistema de interconexión de neuronas que colaboran entre sí para producir un estímulo de salida [22].

La figura siguiente muestra el proceso anteriormente descrito, utilizando una red neuronal siendo implementada en RapidMiner. Dadas las características de la Red, primero se necesita entrenar el modelo, por lo que se particiona la data en 70-30, con lo que se puede testear la calidad de la predicción y dejar preparado el algoritmo para el proceso de producción (implementación en el canal web de venta del SOAP).

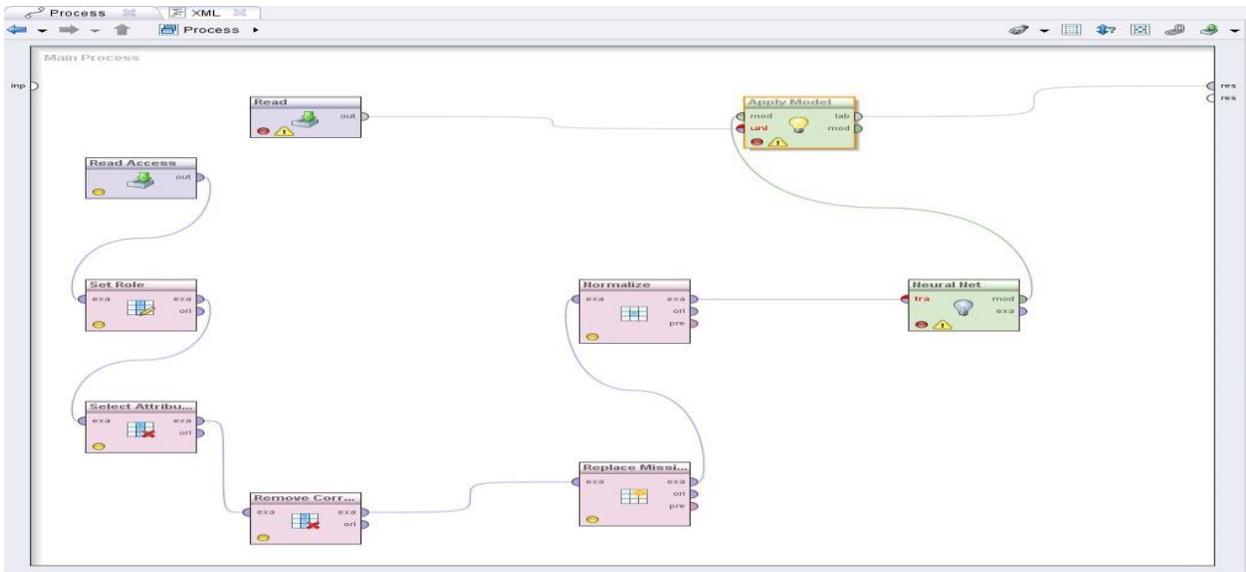


Figura 37: Modelo de Clasificación Red Neuronal RapidMiner. Fuente: Elaboración Propia.

Luego de correr la red neuronal se obtienen los siguientes resultados. Se puede notar que la predicción de los casos “Malos” es considerablemente más baja, pero dado la característica del proyecto esto no es tan relevante, ya que lo principal es poder predecir bien a los clientes “Buenos” ya que estos son candidatos a un descuento.

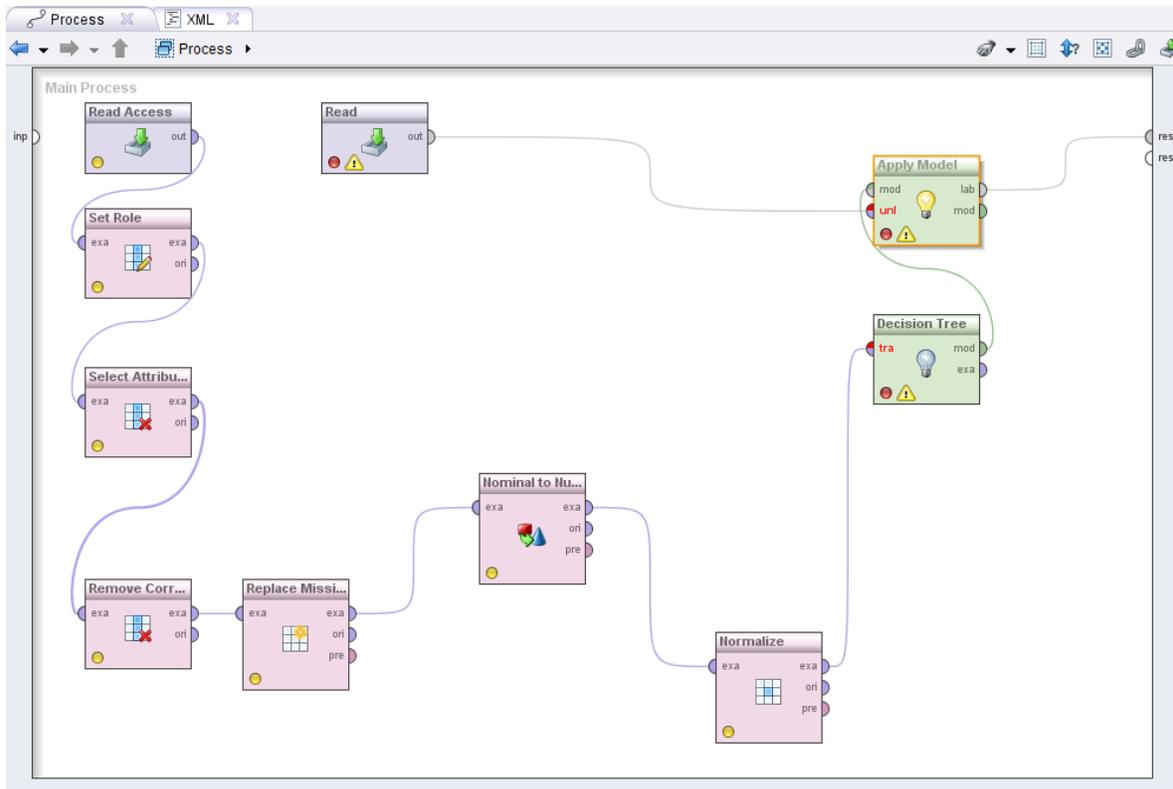
Muestra	Observado	Pronosticado			
		Malo	Riesgoso	Bueno	Porcentaje correcto
Entrenamiento	Malo	2386	550	162	77.0%
	Riesgoso	764	90634	4322	94.7%
	Bueno	5	2324	6323	96.4%
	Porcentaje global	1.9%	56.9%	41.2%	95.1%
Prueba	Malo	294	43	41	77.8%
	Riesgoso	243	9563	1450	85.0%
	Bueno	1	200	6429	97.0%
	Porcentaje global	2.9%	53.7%	43.4%	89.2%

Tabla 12: Resumen Red Neuronal. Fuente: Elaboración Propia.

### 5.1.4.3. Árbol de Decisión

Un árbol de decisión es un modelo de predicción utilizado en el ámbito de la inteligencia artificial. Dada una base de datos se construyen diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que ocurren de forma sucesiva, para la resolución de un problema [22].

Al igual que en el caso anterior se modela el algoritmo escogido en RapidMiner (figura 35). Este algoritmo también debe ser entrenado para poder calibrar de mejor forma sus parámetros relevantes (número de hojas, nodos, etc.).



**Figura 38: Modelo de Clasificación Árbol de Decisión RapidMiner. Fuente: Elaboración Propia.**

En la siguiente tabla se puede observar el resultado del árbol de decisión. Debido a estos resultados y a que el tiempo de procesamiento es similar, es que se escogen las redes neuronales como el algoritmo para clasificar los posibles clientes del SOAP.

Muestra	Observado	Pronosticado			
		Malo	Riesgoso	Bueno	Porcentaje correcto
Entrenamiento	Malo	2128	808	162	68.7%
	Riesgoso	1264	80134	14322	83.7%
	Bueno	2125	5324	58111	88.6%
	Porcentaje global	3.4%	52.5%	44.2%	85.4%
Contraste	Malo	186	110	48	54.2%
	Riesgoso	640	7904	2091	74.3%
	Bueno	436	892	5957	81.8%
	Porcentaje global	6.9%	48.8%	44.3%	76.9%

**Tabla 13: Resumen Árbol de Decisión. Fuente: Elaboración Propia.**

Al comparar las matrices de confusión correspondientes a la red neural y al árbol de decisión se puede apreciar un rendimiento superior en las redes neuronales con un 89% de acierto global contra un 77% del árbol de decisión. Aunque por otro lado el porcentaje de acierto de los casos “buenos” (candidatos a descuento), las redes neuronales estuvieron del orden del 97% mientras que la el árbol de decisión tuvo un 82% de acierto.

A continuación se pueden apreciar las tablas de sensibilidad y especificidad asociadas a cada uno de los métodos de clasificación analizados anteriormente.

La sensibilidad nos indica la capacidad de nuestro estimador para dar como casos positivos los casos que realmente presentan la condición analizada, mientras que la especificidad nos indica la capacidad de nuestro estimador para dar como casos negativos los casos realmente no presentan la condición analizada.

Data	Medida	Clase		
		Malo	Riesgoso	Bueno
Entrenamiento	Sensibilidad	75.6%	96.9%	93.4%
Entrenamiento	Especificidad	99.6%	95.8%	97.6%
Prueba	Sensibilidad	54.6%	97.5%	81.2%
Prueba	Especificidad	99.5%	96.5%	98.1%

**Tabla 14: Sensibilidad y Especificidad Red Neuronal. Fuente: Elaboración Propia.**

Data	Medida	Clase		
		Malo	Riesgoso	Bueno
Entrenamiento	Sensibilidad	38.6%	92.9%	80.0%
Entrenamiento	Especificidad	99.4%	91.1%	91.9%
Prueba	Sensibilidad	14.7%	88.7%	73.6%
Prueba	Especificidad	99.1%	86.9%	86.9%

**Tabla 15: Sensibilidad y Especificidad Árbol de Decisión. Fuente: Elaboración Propia.**

El porcentaje de acierto que arrojan las matrices de confusión, sumado a la sensibilidad y especificidad de cada modelo refleja una diferencia considerable, en especial cuando se tiene en cuenta la intención de premiar a los “Buenos” candidatos y no recargar a los otros, por lo que se decidió por implementar la red neuronal antes que el árbol de decisión, a pesar de que mantener una red neural en general es más costoso que otras opciones.

## Capítulo 6

# Arquitectura TI

### 6.1. Diagramas UML

El Lenguaje Unificado de Modelado (UML) es el lenguaje de modelado de sistemas de software más empleado en la actualidad. Es un lenguaje gráfico para visualizar, especificar, construir y documentar un sistema. UML ofrece un estándar para describir un "plano" del sistema (modelo), incluyendo aspectos conceptuales tales como procesos de negocio, funciones del sistema, y aspectos concretos como expresiones de lenguajes de programación, esquemas de bases de datos y compuestos reciclados. UML se utiliza para definir un sistema, para detallar los artefactos en el sistema y para documentar y construir. En otras palabras, es el lenguaje en el que está descrito el modelo.

Para poder visualizar los sistemas TI que deber ser construidos conforme completar el presente proyecto, es que se presentan algunos diagramas UML, tales como; Diagrama de paquetes, Casos de usos, Diagrama de Clases, Diagrama de secuencia y diagrama de despliegue. Estos diagramas pretenden presentar de forma simple y concisa el funcionamiento de la aplicación deseada y facilitarán la implementación final por el equipo de tecnología de la empresa [10].

### 6.1.1. Diagrama de Paquetes

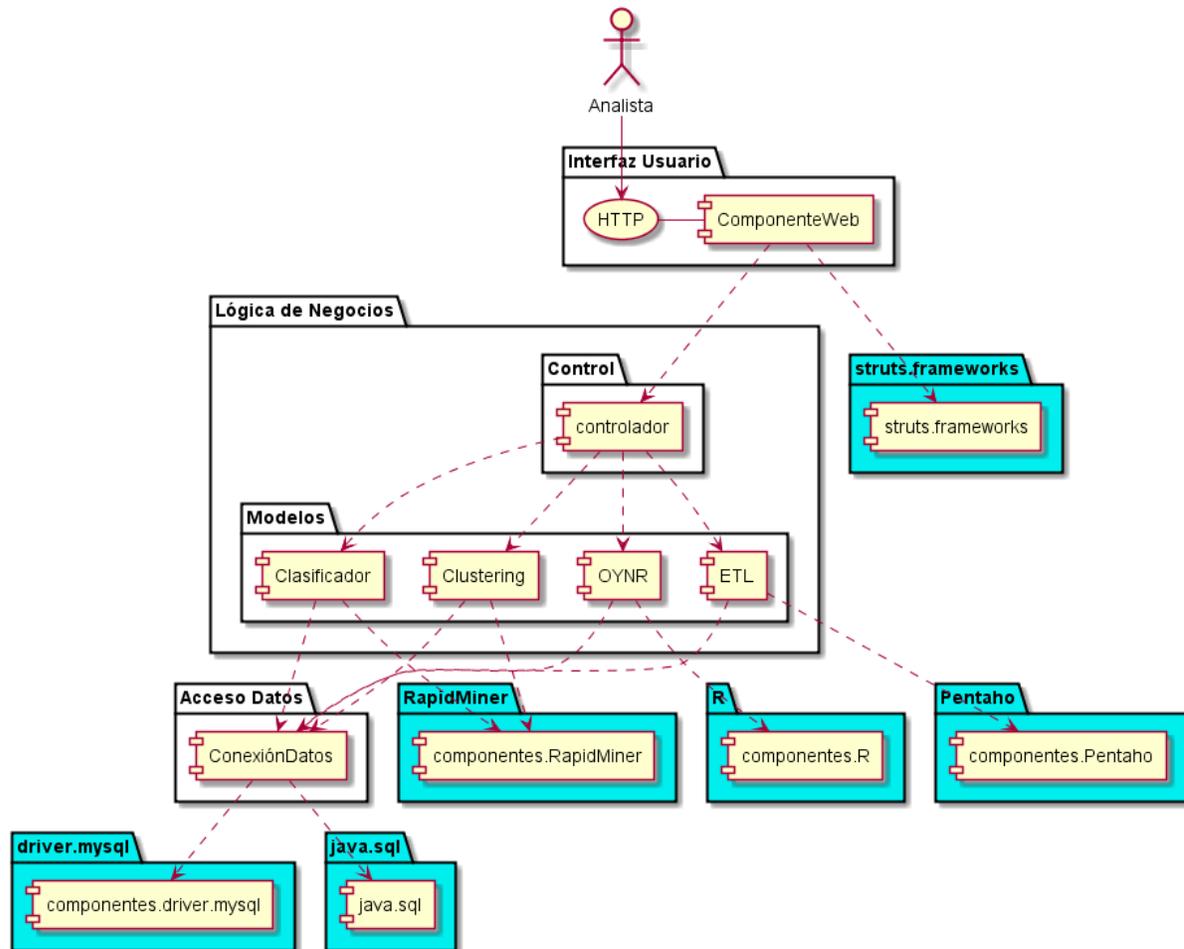


Figura 39: Diagrama de Paquetes. Fuente: Elaboración propia.

El diagrama de paquetes anterior, muestra que paquetes son necesarios para el funcionamiento de la aplicación, en este caso gran parte de la aplicación se la lleva la preparación de los datos, para luego ejecutar los modelos apropiados y poder obtener resultados. Evidentemente para poder lograr lo anterior, se debe tener conexión a las bases de datos adecuadas.

### 6.1.2. Casos de Uso

Los casos de uso son una descripción de los pasos o las actividades que deberán realizarse para llevar a cabo algún proceso. En el caso del proyecto, el caso de uso representa las actividades que realizara el usuario para poder completar el proceso descrito en los capítulos anteriores.

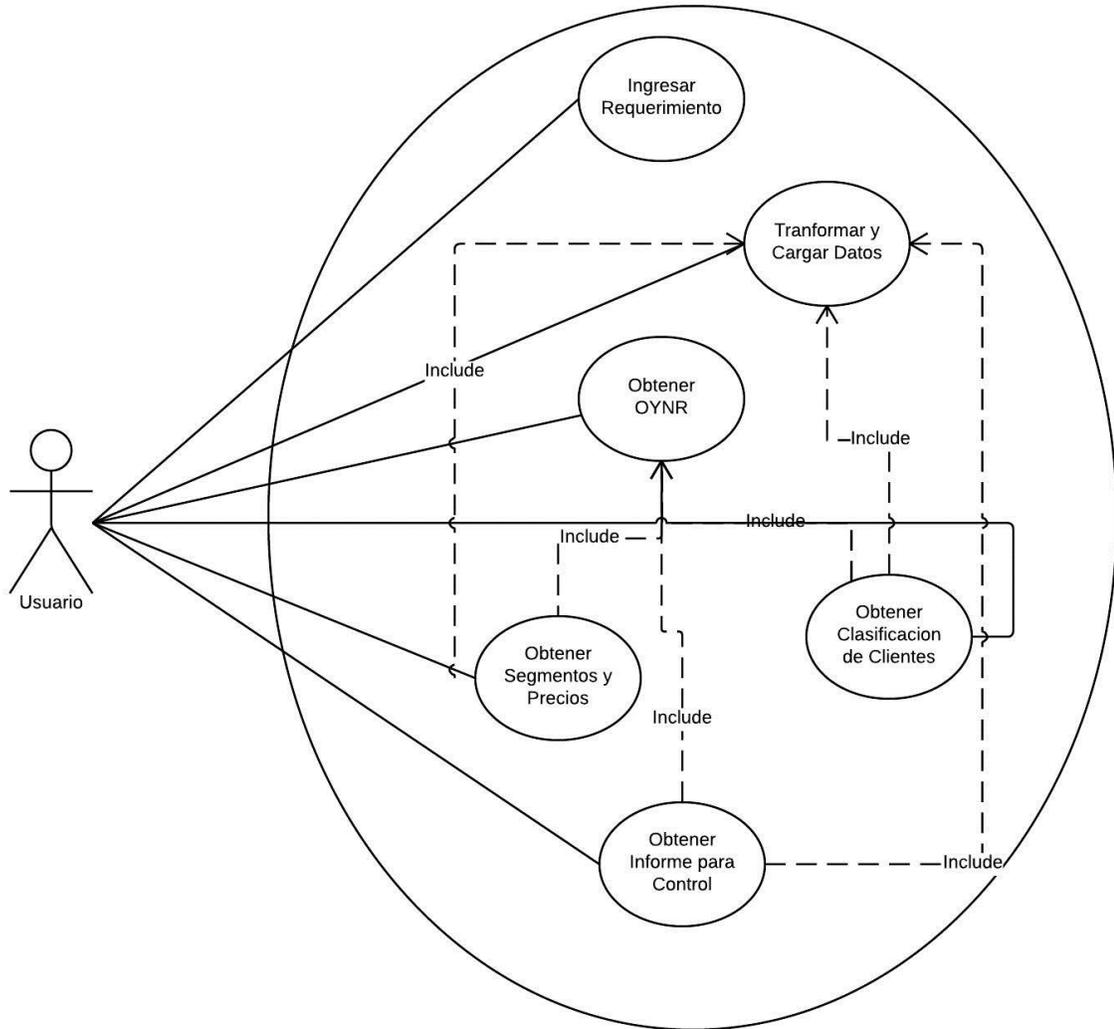
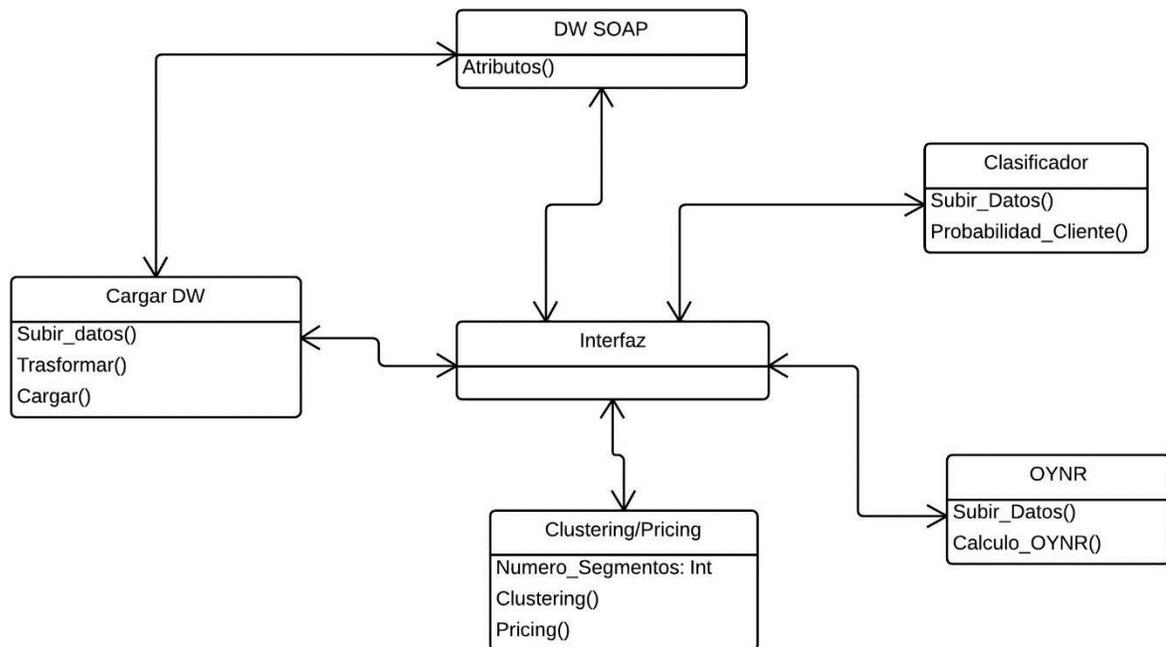


Figura 40: Caso de Uso. Fuente: Elaboración propia.

### 6.1.3. Diagrama de Clases

Un diagrama de clases, es un tipo de diagrama estático que describe al estructura de un sistema mostrando las clases del mismo, sus atributos, operaciones, métodos y la relación entre ellos. En el caso del proyecto las clases son la encargadas de realizar las actividades principales del sistema, como realizar cálculos, cargar bases, etc.



**Figura 41: Diagrama de clases. Fuente: Elaboración propia.**

#### 6.1.4. Diagrama de secuencia

Un diagrama de secuencia es un diagrama UML de interacción que muestra cómo operan los procesos entre sí y en qué orden. Es una construcción de un gráfico de secuencia de mensajes. Un diagrama de secuencia muestra las interacciones de objetos dispuestos en secuencia de tiempo.

Representa los objetos y clases que participan en el escenario y la secuencia de mensajes intercambiados entre los objetos necesarios para llevar a cabo la funcionalidad del escenario deseado. Los diagramas de secuencia son típicamente asociados con las realizaciones de casos de uso en la vista lógica del sistema en desarrollo.

El caso de uso mostrado anteriormente queda representado como diagrama de secuencia en la siguiente figura, donde se puede ver cómo interactúan las clases ya definidas y como el usuario puede realizar todo el proceso diseñado en los capítulos anteriores.

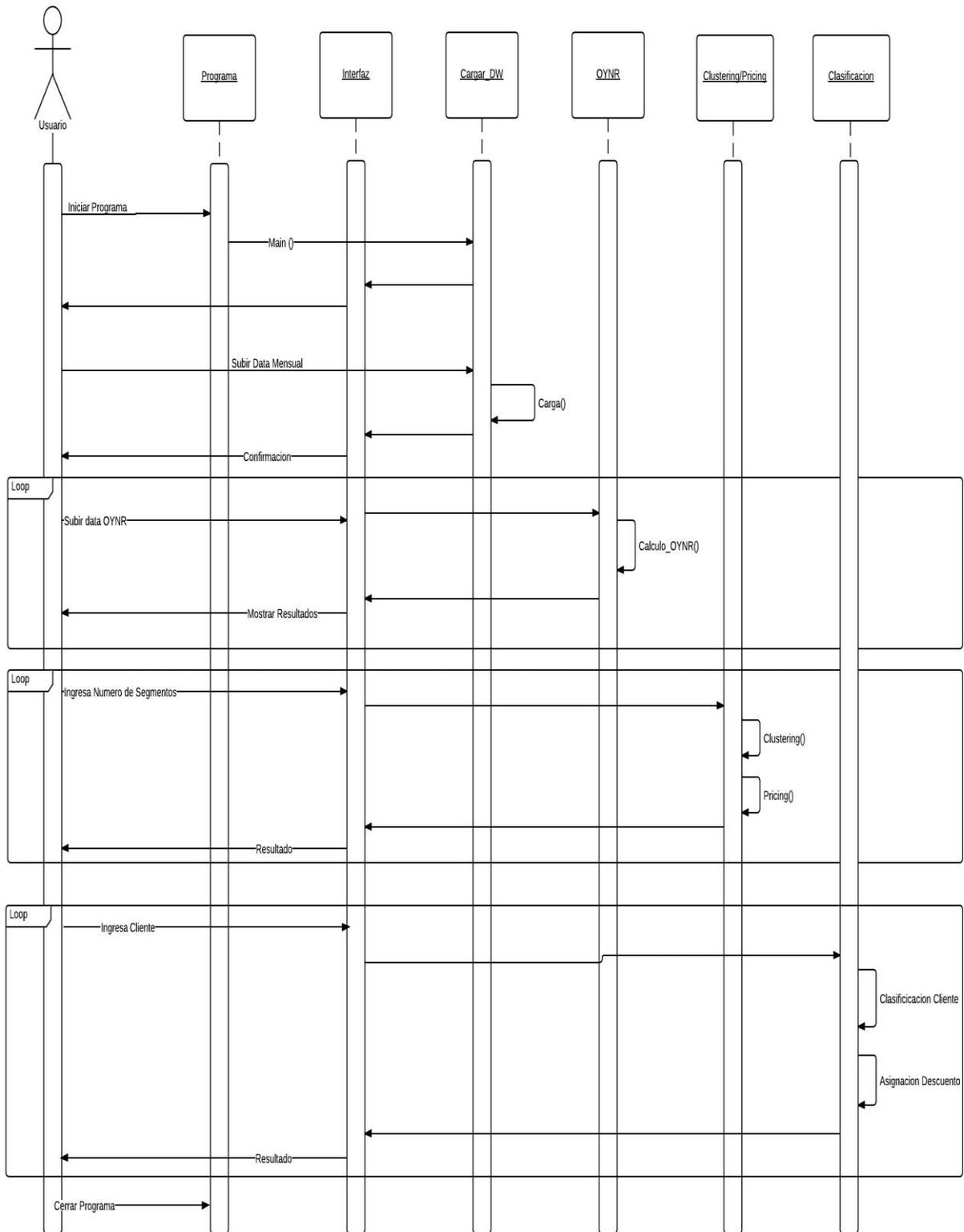


Figura 42: Diagrama de Secuencia. Fuente: Elaboración propia.

### 6.1.5. Diagrama de despliegue

El Diagrama de Despliegue es un tipo de diagrama UML que se utiliza para modelar la disposición física de los artefactos software en nodos. Es útil para visualizar que componente de software corre sobre que nodo y como las distintas piezas se conectan entre sí. A continuación se aprecia el diagrama de despliegue del proyecto, haciendo énfasis en los componentes principales para su funcionamiento; R, Rapidminer y Pentaho.

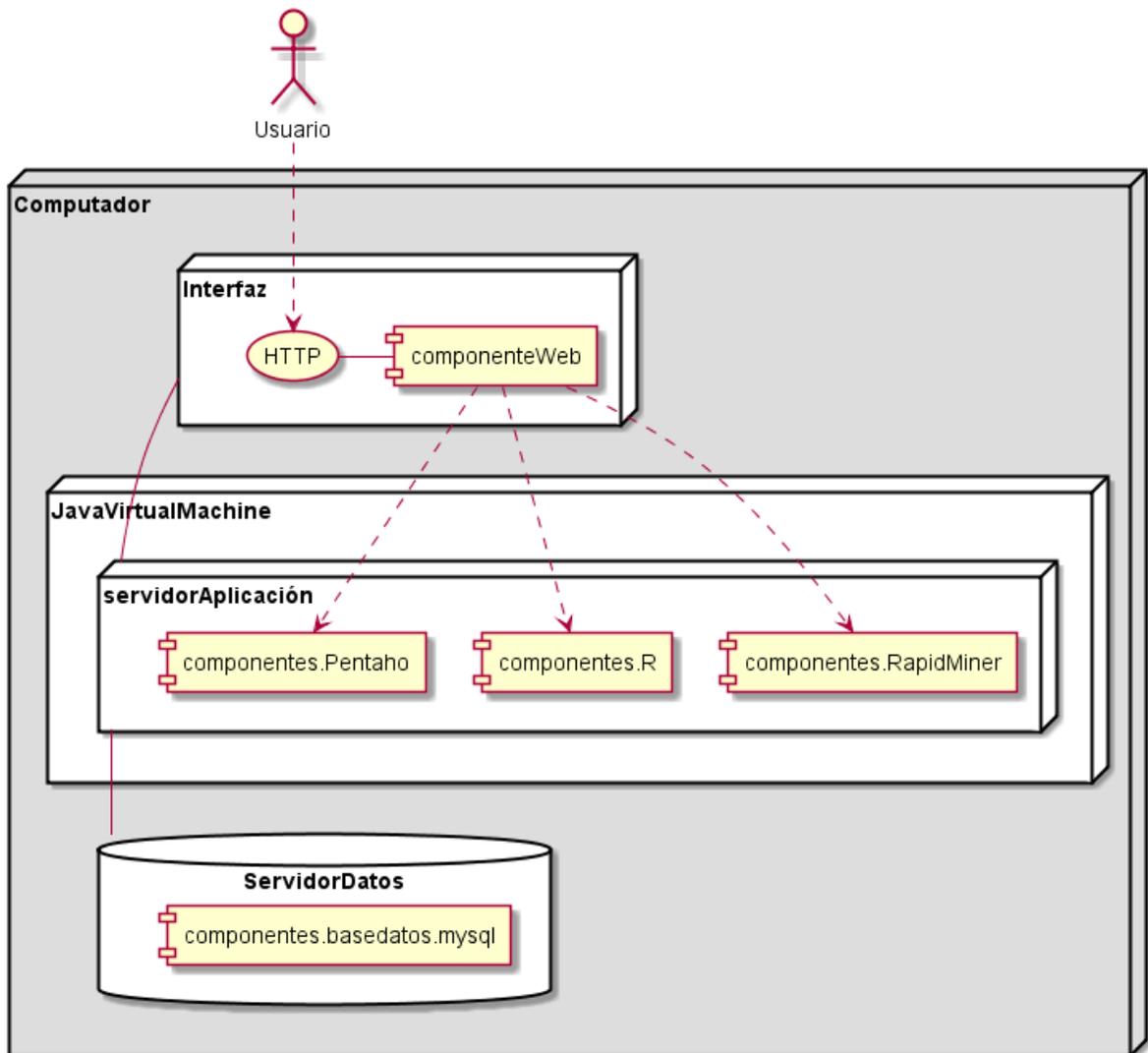


Figura 43: Diagrama de Despliegue. Fuente: Elaboración propia.

## Capítulo 7

### Resultados

#### 7.1 Resultados Segmentación

La segmentación presentada a continuación corresponde al análisis realizado en el capítulo 5.1.3 en donde se terminó escogiendo como algoritmo de Clustering a K-means, se determinó la utilización de 5 segmentos y finalmente se realizó una validación interna correspondiente al cálculo del coeficiente Silhouette, el cual entregó un valor de 0.843, donde el máximo es de 1, para una segmentación perfecta.

La siguiente tabla muestra el costo medio de la muerte y los gastos médicos (en UFs), así como la frecuencia y el número de afectados de las mismas coberturas con respecto a los segmentos encontrados.

Cluster	CM Gastos Medicos	CM Muerte	Frec Gastos Medicos	Frec Muerte	N Afectados GM	N Afectados M
cluster_0	29	330	11	0,52	1,35	1,13
cluster_1	23	327	13	0,57	1,42	1,13
cluster_2	28	288	14	0,63	1,30	1,01
cluster_3	24	308	15	0,73	1,35	1,07
cluster_4	15	299	24	0,31	1,37	1,00
<b>Total general</b>	<b>25</b>	<b>312</b>	<b>14</b>	<b>0,66</b>	<b>1,36</b>	<b>1,08</b>

**Tabla 16: Características Clusters. Fuente: Elaboración Propia.**

La tabla anterior muestra la distribución geográfica de cada uno de los clusters y el número de asegurados en cada uno de ellos.

Cluster	REGION	Numero Asegurados
cluster_0	Metropolitana	280441
cluster_1	Antofagasta	64435
	Arica y Parinacota	2428
	Atacama	2641
	Coquimbo	23921
	Magallanes	42167
	O'Higgins	24870
	Tarapacá	38500
	Valparaíso	49697
cluster_2	Metropolitana	265403
cluster_3	Metropolitana	786855
cluster_4	Araucanía	12697
	Bío Bío	10748
	Los Lagos	20664
	Los Ríos	3741
	Maule	10372
<b>Total general</b>		<b>1639580</b>

**Tabla 17: Distribución geográfica Clusters. Fuente: Elaboración Propia.**

En las siguientes tablas de puede observar la antigüedad promedio de los vehículos asegurados y la siniestralidad inventario de las principales marcas. Con respecto a las marcas es muy interesante notar que marcas como el BMW tiene una siniestralidad tan baja, que a priori se podría pensar que se le debería asignar una prima menor a la que se le ha asignado en el pasado. Esta información será particularmente útil cuando se intente clasificar los clientes para otorgar algún tipo de descuentos.

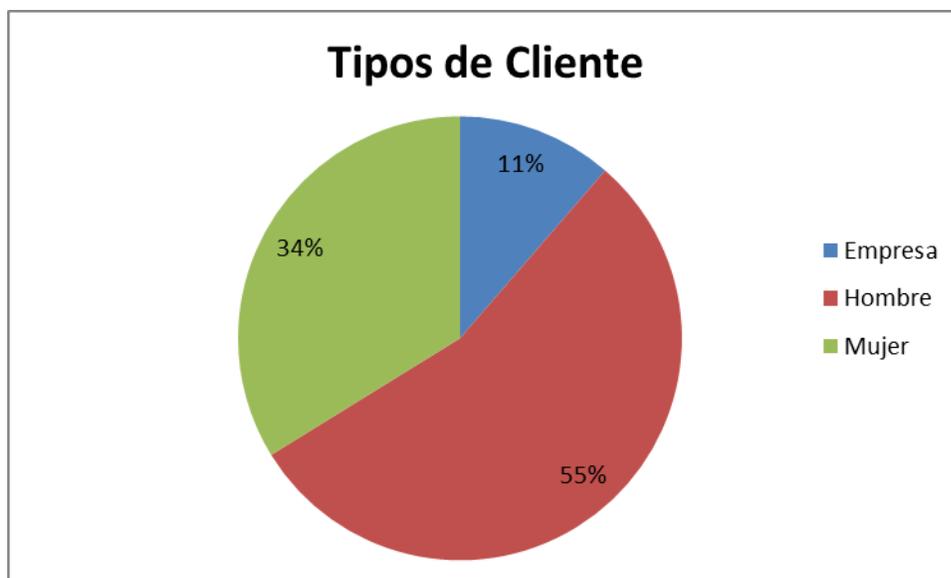
Cluster	Antigüedad Vehículo Promedio [Años]
cluster_0	17,49
cluster_1	2,93
cluster_2	11,22
cluster_3	4,03
cluster_4	20,06

**Tabla 18: Antigüedad Promedio vehículos Clusters. Fuente: Elaboración Propia.**

Marca	Siniestralidad Inventario	Numero de Polizas
BMW	26%	9.121
CHEVROLET	68%	31.562
CITROEN	88%	14.182
DAEWOO	82%	6.413
DAIHATSU	141%	11.462
DODGE	37%	14.096
FIAT	75%	17.895
FORD	80%	24.727
HONDA	64%	15.867
HYUNDAI	115%	32.729
JEEP	126%	11.101
KIA MOTORS	68%	21.112
MAZDA	69%	19.693
MERCEDES BEN.	51%	12.072
MITSUBISHI	109%	26.571
NISSAN	94%	26.714
OPEL	32%	6.012
PEUGEOT	89%	19.845
RENAULT	89%	14.188
SAMSUNG	66%	5.045
SSANGYONG	153%	10.511
SUBARU	121%	14.475
SUZUKI	90%	30.084
TOYOTA	94%	29.418
VOLKSWAGEN	106%	16.572
<b>Total general</b>	<b>87%</b>	<b>441.467</b>

**Tabla 19: Siniestralidad Marcas más representativas. Fuente: Elaboración Propia.**

Finalmente se pueden ver las diferencias entre los distintos tipos de clientes, en donde se clasifican en hombres, mujeres y empresas. Se pueden observar ciertas diferencias pero como se había adelantado en el primer análisis de las variables discriminantes, el sexo no resulta ser tan relevante como se podría imaginar.



**Figura 44: Tipos de Clientes SOAP. Fuente: Elaboración Propia.**

Tipo	Siniestralidad
Empresa	62%
Hombre	47%
Mujer	46%
<b>Total general</b>	<b>49%</b>

**Tabla 20: Distribución Tipo de Clientes. Fuente: Elaboración Propia.**

La tabla anterior representa la siniestralidad correspondiente a cada tipo de clientes. De forma, quizás, sorprendente no existe diferencia estadística entre los hombres y las mujeres, pero si se demuestra claramente que los seguros comprados por las empresas resultan ser más siniestrosos.

Tipo	CM_Gastos_Medicos	CM_Muerte	Frec_Gastos_Medicos	Frec_Muerte	Numero de Polizas
Empresa	32,8	317	16,9	1,00	186.173
Hombre	28,8	305	13,9	0,54	899.141
Mujer	23,7	327	15,1	0,48	554.231
<b>Total genera</b>	<b>27,5</b>	<b>314</b>	<b>14,6</b>	<b>0,57</b>	<b>1.639.545</b>

**Tabla 21: Características Tipos de Clientes. Fuente: Elaboración Propia.**

Para entender mejor cómo se comportan los distintos tipos de clientes, se verifico sus costos medios y frecuencias de siniestros. De aquí se obtiene otra información interesante, y esta es que las mujeres tienen gastos médicos menores pero su costo de muertes es mayor, porque finalmente las mujeres tienden a ir con más personas en sus vehículos (en comparación con las empresas y los hombres).

Los resultados del Clustering muestran varios resultados interesantes y permiten tener una mejor segmentación de la cartera y a la vez entender mejor el perfil de los asegurados de Consorcio.

Esta información es particularmente relevante, ya que a partir de esta segmentación, es que calculara una prima técnica indicativa (el departamento de vehículos es el encargado de calcular esta prima), y los informes de gestión correspondientes harán énfasis en el comportamiento de los segmentos, lo que a la larga también debería afectar como se estimaran los presupuestos en los próximos años y cuáles serán las mejores decisiones comerciales a tomar.

## 7.2. Pricing

Con respecto al pricing, vale destacar que lo que se calculara no es el “precio” final, si no el “precio técnico” del seguro, que corresponde a la prima inventario (o prima de riesgo). La prima total que se calcula para cada cliente se puede aproximar de la siguiente forma:

$$\text{Prima} = \text{Prima Inventario} + \text{Gastos} + \text{Comision} + \text{Margen}$$

Como se puede apreciar la prima inventario es el único parámetro modificable, ya que, los demás corresponden a decisiones de otros niveles (directores) o valores que no se pueden modificar fácilmente como la comisión.

Actualmente se tienen un modelo en Excel que calcula la prima inventario teórica por segmento. Y lo que se planea hacer, es encontrar los segmentos adecuados (paso anterior) para luego calcular la prima inventario teórica en cada periodo (se tienen 5 años de historia) y de con esos datos proyectar la del periodo siguiente.

$$\text{Prima Inventario} = \text{Frec}_{Global} * \text{Prob}_{XX} * N_{Afectados_{XX}} * \frac{CM_{XX_{Rut}}}{1000000} * 365$$

En donde,

$\text{Frec}_{Global}$  = Es la frecuencia con que ocurre los siniestros

$\text{Prob}_{XX}$  = Es la probabilidad de que se use la cobertura “XX” dado un siniestro

$N_{Afectados_{XX}}$  = Numero de accidentados con XX por siniestro con XX

$CM_{XX_{Rut}}$  = Costo Medio XX a nivel de accidentados

XX= {Gastos Médicos, Muerte, Invalidez Total, Invalidez Parcial}

La fórmula anterior, es solo una modificación de la formula clásica para el cálculo de la prima de riesgo [22].

$$\text{Prima Inventario} = \text{Frecuencia} * \text{Severidad} * \text{Monto Asegurado}$$

Esta fórmula se debe cumplir para cualquier seguro. En general la frecuencia y la severidad (o costo medio) vienen dados por funciones de probabilidad adecuadas a cada cartera (por ejemplo; una distribución Poisson y una normal respectivamente).

Debido a la utilización de esta fórmula es que encontrar los segmentos apropiados es tan fundamental. El análisis probabilísticos y las proyecciones de los segmentos quedan en manos del departamento de vehículos que se encarga de la tarificación, pero como parte del proyecto se definió la entrega de la prima teórica al momento de cálculo.

Región	Prima Total [UF]	Prima Inventario [UF]	Prima Ganada [UF]	Exposición Mensual	Número de Pólizas	Número de Siniestros	Monto Siniestros [UF]	Monto OYNR [UF]	Costo Medio [UF]
<b>IV</b>									
Metropolitana	14,776	0.2338	142,199	4,117,411	23,973	1,506	66,173	13,951	42.19
Tarapacá	176	0.2209	6,560	207,750	457	52	3,315	513	60.73
Antofagasta	1,651	0.1395	6,657	161,286	2,393	37	1,561	308	39.56
Valparaíso	26	0.1245	4,535	135,879	85	57	1,104	306	20.36
Magallanes	960	0.1402	4,154	101,918	1,559	40	1,015	170	21.79
O'Higgins	282	0.2235	3,126	85,191	745	51	1,320	261	23.98
Coquimbo	34	0.1041	2,513	65,388	79	17	460	107	24.12
Los Lagos	1	0.1473	1,576	34,179	1	13	354	66	23.01
Araucanía	1,174	0.0964	2,044	31,313	1,462	23	195	51	8.11
Maule	121	0.2422	620	14,009	269	10	253	28	24.73
Bío Bío	0	0.0530	430	10,538	0	4	34	13	9.35
<b>Total general</b>	<b>19,201</b>	<b>0</b>	<b>174,413</b>	<b>4,964,862</b>	<b>31,023</b>	<b>1,810</b>	<b>75,784</b>	<b>15,775</b>	<b>40.03</b>

**Tabla 22: Cálculo Prima Inventario. Fuente: Elaboración Propia.**

La prima inventario se controla con el indicador llamado “Siniestralidad Inventario” que es el cociente entre todos los siniestros y la prima inventario, idealmente este indicador debe ser 100%, es decir, los siniestros deberían ser iguales a la prima asignada para siniestros.

Lógicamente cuando el indicador supera el 100% (los siniestros fueron mayores a la prima inventario), es un problema, ya que implica, que ese porcentaje extra debe salir de otro lado, como por ejemplo el margen o simplemente generar pérdidas.

A priori se podría pensar que siniestralidades bajas son ideales, dado que aumentan inmediatamente el margen percibido, pero esto no es así, debido a que la política de la compañía tiene como motivación que la siniestralidad inventario sea 100%, porque siendo menor se asume que se pudo haber cobrado un precio menor y por lo tanto vender más pólizas y obtener un mayor marketshare. Esta idea se basa en la creencia de una elasticidad positiva (un mercado elástico). Este tema es analizado con mayor profundidad en el siguiente capítulo.

Con todos estos antecedentes ya se puede crear el informe final que utilizaran las áreas de vehículo y de SOAP. Para vehículos se entregan todas las primas teóricas por segmentos y sub-segmentos (además de resultados contables), mientras que para SOAP (área comercial) se hace énfasis en las variables más relevantes que explican los resultados, para que ellos tomen medidas y puedan atacar mejor los targets escogidos.

Estos informes serán entregados de forma mensual. Para lograr esto es que la automatización se vuelve primordial.

### 7.3. Prototipo

Todos los procesos anteriores deben consolidarse en un programa que utilice java como lenguaje orquestador. La siguiente figura muestra la interfaz del prototipo creado, es capaz de validar la data, calcular el OYNR y realizar el Clustering.

El prototipo es capaz de ejecutar los procesos creados tanto en RapidMiner como en Pentaho. La clasificación es entrenada dentro de los procesos creados, pero su ejecución es parte del prototipo solo de forma temporal ya que su implementación funcional debe ser dentro del sistema web de Consorcio y esta será encargada por el área TI de la empresa.

En la siguiente figura se puede apreciar la interfaz general del prototipo, en donde se puede escoger entre los distintos procesos creados anteriormente. Los procesos deben ejecutarse en orden, si por ejemplo se ejecuta el informe final antes de haber ejecutado los demás procesos el programa se cae al no encontrar los datos correspondientes.

Cada botón (sin contar el informe final) lleva a una interfaz similar donde se piden algunos datos (e.g.: dirección de los datos a utilizar) y la posibilidad de ejecutar el proceso. Parámetros más específicos de algunos procesos solo pueden ser modificados dentro de los programas que corren internamente (Pentaho, RapidMiner, etc.), esto debe ser modificado en la versión final.

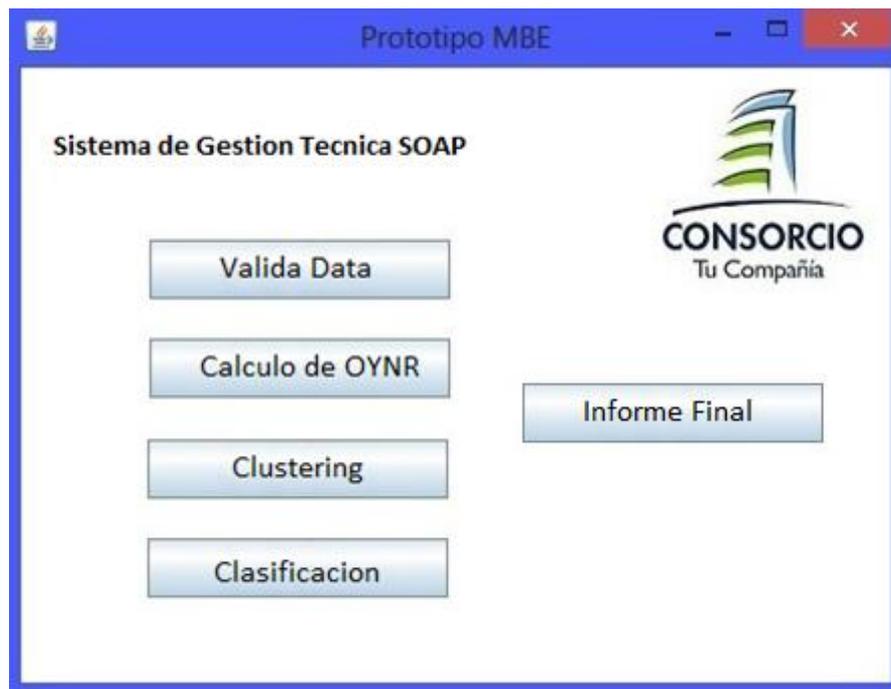


Figura 45: Interfaz Prototipo. Fuente: Elaboración Propia.



Figura 46: Interfaz Prototipo/Validación Data. Fuente: Elaboración Propia.



Figura 47: Interfaz Prototipo/OYNR. Fuente: Elaboración Propia.



Figura 48: Interfaz Prototipo/Clustering. Fuente: Elaboración Propia.



Figura 49: Interfaz Prototipo/Clasificación. Fuente: Elaboración Propia.

El sistema TI final será implementado por completo por el área de proyectos y el área TI de la empresa, como entregables se consideran todos los modelos en los diferentes programas utilizados (RapidMiner, Pentaho, R), los diagramas UML y el prototipo capaz de ejecutar los procesos antes descritos.

La idea final es que exista una aplicación de escritorio para la gestión técnica del SOAP (ETL, Clustering, generación de informes, etc.) y una modificación para el servicio web actual de venta de SOAP.

## 7.4. Gestión del Cambio

Cualquier proyecto de cambio necesita de la gestión del cambio para poder materializarse. La literatura nos muestra muchos casos de proyectos de todos los tamaños que han fallado (a pesar de contar con personas y recursos idóneos) al no considerar la gestión del cambio como un insumo fundamental para el proyecto [25].

En la siguiente figura se puede observar un modelo para la gestión del cambio [26], en donde se deben cumplir una serie de etapas para asegurar que el proyecto de cambio sea implementado con exitoso. Además existen un par de procesos que son transversales a todo el proyecto como la comunicación entre todos los actores relevantes y la gestión de los estados de ánimos de los equipos participantes.



Figura 50: Modelo de Gestión de Cambio. Fuente: Elaboración Propia.

#### **7.4.1. Análisis Situación Actual**

En esta etapa se analizó la situación actual de la empresa, la estrategia de la misma y las proyecciones futuras para el producto escogido. Con esta información se concluyó que se recitaba un cambio, el que queda plasmado dentro los alcances del presente proyecto.

#### **7.4.2. Formación de Equipos**

El siguiente paso fue el de crear equipos capaces de liderar, promocionar, diseñar e implementar el proyecto definido previamente. Se definió al líder de proyecto (Gonzalo Araya) y los departamentos que prestarían funciones de apoyos por su mayor conocimiento del producto (Departamento de vehículos y departamento de SOAP).

#### **7.4.3. Crear y comunicar visión**

Conforme aumentar la probabilidad de éxito del proyecto es que fue necesario crear una visión propia, que hiciera énfasis en la necesidad del proyecto y como este produciría beneficios reales para Consorcio y sus empleados. Además de crear (y validar) la visión, fue necesario comunicarla a todas las personas que de una u otra firma serían afectadas por este proyecto. De esta forma hubo un par de semana caracterizada por reuniones informativas para comunicar la nueva visión del proyecto.

#### **7.4.4. Avanzar por etapas**

Para evitar el riesgo de que el proyecto pareciera no avanzar o no tener hitos positivos, es que se definió una estrategia de avance por etapas, en donde siempre se definía una etapa (u objetivo) y luego se comunicaban los avances y los términos de las etapas para poder pasar a la siguiente. Entre las etapas destacadas se pueden nombrar el diseño de los nuevos procesos, los resultados de los distintos algoritmos y finalmente los resultados iniciales del prototipo.

#### **7.4.5. Desarrollar Habilidades**

En general los proyectos de cambio requieren del desarrollo de nuevas habilidades para muchos de los actores (en especial por que estos proyectos incluyen aprendizajes de segundo grado). A pesar de esto, por la naturaleza del proyecto no requirió del desarrollo de grandes habilidades, ya que los desarrollos finales resultaron en un prototipo bastante sencillo de utilizar por todos los involucrados.

#### **7.4.6. Evaluación y Cierre**

En el desarrollo de proyectos siempre es necesario tener un cierre del mismo que incluya una evaluación y destaque los logros, los fracasos y los aprendizajes asociados. En el caso del proyecto la evaluación fue positiva, ya que se cumplieron los objetivos del mismo y no se sobrepasaron de manera considerable los recursos destinados (se atrasaron algunas fechas de entregas, pero por razones de fuerza mayor). El cierre incluyó la puesta en marcha del prototipo (de forma oficial), que sigue siendo utilizado hasta el día de hoy.

#### 7.4.7. Roles y estructura organizacional

Un punto muy relevante en la gestión del cambio es la definición de roles y la estructura organizacional del proyecto.

En la figura siguiente se puede apreciar la estructura de roles del proyecto. En este caso se tiene como cliente principal al departamento de vehículos (los encargados de la tarificación), se utiliza el apoyo experto del departamento de SOAP (por ejemplo en temas normativos y de comercialización del producto), además de la necesidad del marketing interno necesario para mantener sano el proyecto durante su realización.

Finalmente se define al programador y al diseñador web, como roles dependientes del líder de proyectos, debido a que están encargados de dar los toques finales al proyecto, llevando el prototipo a su versión definitiva.

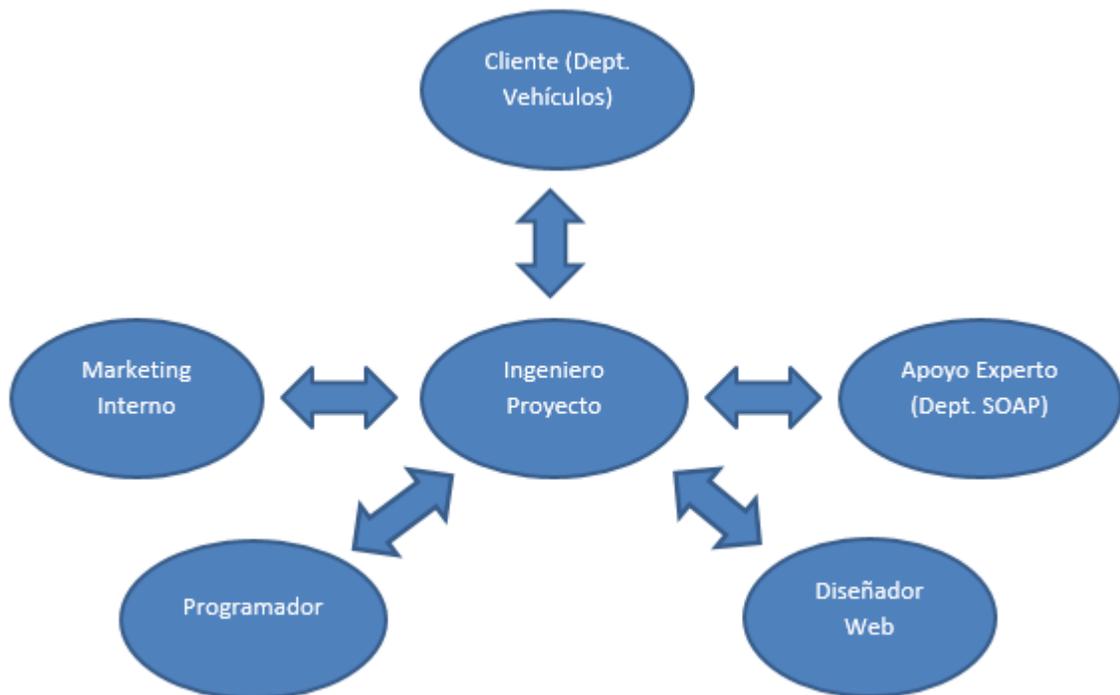


Figura 51: Mapa de Roles proyecto. Fuente: Elaboración Propia.

## 7.5. Evaluación Implementación

Desde el mes de septiembre el proyecto fue implementado a cabalidad, reemplazando los procesos anteriores en la gestión técnica del SOAP. Debido a las fechas (ya fueron ejecutadas todas las campañas del 2014), los resultados de la analítica no han sido utilizados aun para mejorar la tarificación, pero los informes si han sido utilizados para el control del producto. Un ejemplo claro de los beneficios del proyecto es que comparado con los meses anteriores el tiempo requerido para generar el informe de siniestralidad del SOAP paso de demorarse una par de días a demorarse una tarde, además de incluir información mucho más rica que en ocasiones anteriores. Por primera vez las distintas áreas han podido ver los resultados del SOAP con vistas asociadas al sexo, la marca del vehículo o su antigüedad. Esto ha llevado a modificar algunos parámetros relevantes en el presupuesto de SOAP para el año 2015.

Con respecto a la implementación final, el área de proyectos de la empresa se encuentra con atrasos en varios proyectos estratégicos de la compañía (Reservas y Reaseguro), por lo que la implementación en su forma definitiva del proyecto ha quedado pendiente hasta nuevo aviso. A pesar de esto, el prototipo no ha presentado ningún problema en los procesos realizados en septiembre, octubre, noviembre y diciembre.

## 7.6. Evaluación Económica

### 7.6.1. Estimación Demanda

Dado las características del proyecto es necesario contar la función demanda del SOAP, para de esta forma poder calcular la elasticidad de la misma y así poder estimar cuáles serán los beneficios potenciales de la reducción de precio que se pretende con la mejor tarificación y los descuentos selectivos.

Para estimar la demanda, se utilizó un enfoque clásico al utilizar una regresión lineal con los datos que se disponían. Los datos utilizados para la regresión fueron el precio promedio del seguro en los últimos 5 años (prima comercial), la cantidad de seguros vendidos, la comisión pagada (asumiendo mayor comisión pagada como un esfuerzo de marketing) y el parque automotriz de cada periodo.

Dado lo anterior se generó la siguiente tabla:

Año	Numero de Polizas	Prima Promedio [UF]	Comsion [UF]	Numero de Vehiculos
2009	206.719	0,528	8.494	3.068.220
2010	182.557	0,515	2.354	3.299.446
2011	230.892	0,502	18.684	3.571.219
2012	458.047	0,414	29.733	3.885.581
2013	590.129	0,369	24.250	4.168.980

Tabla 23: Demanda de SOAP. Fuente: Elaboración Propia.

Con lo anterior se calculó la regresión correspondiente, obteniendo los siguientes resultados:

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,999 <sup>a</sup>	,998	,994	,01777	,998	216,798

**Tabla 24: Resumen Regresión Demanda (1). Fuente: Elaboración Propia.**

Modelo	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Regresión	,205	3	,068	216,798	,050 <sup>b</sup>
Residual	,000	1	,000		
Total	,206	4			

**Tabla 25: Resumen Regresión Demanda (2). Fuente: Elaboración Propia.**

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
(Constante)	10,467	3,301		3,171	,194
LPrima	-3,487	,389	-1,051	8,955	,071
LCom	,118	,029	,234	4,145	,151
LNVe	1,014	,527	-,239	1,925	,305

**Tabla 26: Resumen Regresión Demanda (3). Fuente: Elaboración Propia.**

Como queda expresado en las tablas anteriores, la regresión tiene un ajuste muy bien a los datos y estos son muy significativos a la hora de explicar los cambios de la demanda del SOAP.

Además se pueden apreciar los coeficientes respectivos para cada variable y de esta forma comprender como afecta cada uno de estos a la demanda final.

Dada la regresión anterior se pudo calcular la función demanda del SOAP (dejando constantes todas las variables menos el precio) y esta queda plasmada en la siguiente figura.

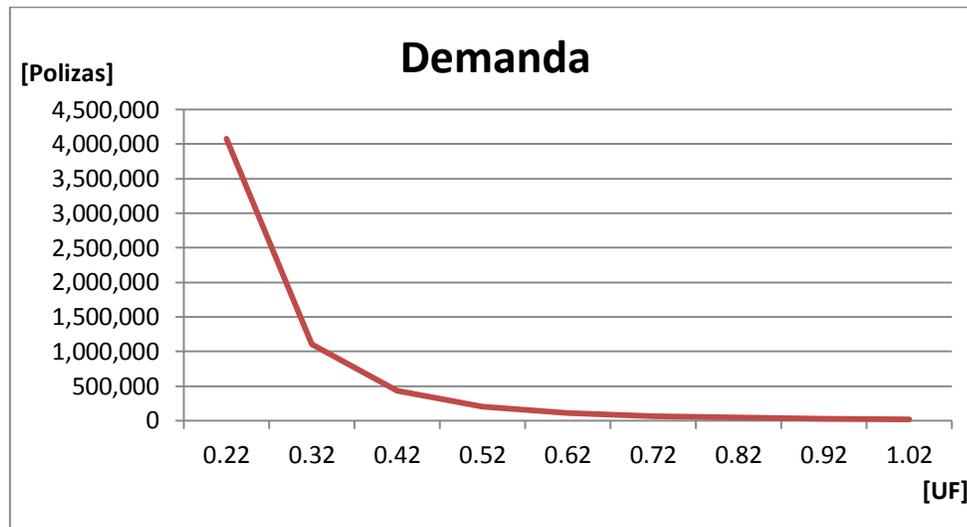


Figura 52: Demanda de SOAP

Con estos resultados se pudo calcular la elasticidad de la demanda. Vale destacar que se utilizó el enfoque de la “elasticidad arco” ya que la elasticidad no es la misma en toda la curva, solo se consideró el “arco” de precios en donde se sabe que podría oscilar el valor del seguro siguiendo siendo rentable.

$$E_a = 2,6$$

Que la elasticidad tenga este valor (en valor absoluto) indica que la demanda es elástica y al reducir el precio el aumento de seguros (a un valor menor) tiene un efecto positivo en el beneficio total.

Algo interesante del gráfico de demanda, es que si se cobrara la mitad de lo que hoy se cobra (valor similar a las demás compañías competidoras), consorcio podría atraer a todo el parque automotriz, esto parece tener mucho sentido, si pensamos en que la mayor diferencia entre los SOAP es su precio. El problema radica, obviamente, que dado todos los siniestros esperados el negocio tendría un margen muy negativo (a pesar de absorber todo el market share).

### 7.6.2. Inversión Inicial

La inversión inicial es bastante modesta, principalmente dado que para los modelos se utiliza software libre, el computador donde se correrán existe y tiene tiempo ocioso (ya que fue adquirido para correr procesos), por lo tanto los únicos costos que se pueden considerar como inversión inicial, son las horas hombre del ingeniero líder del proyecto (al diseñar y trabajar en el proyecto) y el trabajo del programador, que se encargara de diseñar las versión final del sistema TI descrito.

No se consideran costos de capacitación, ya que esta es casi inexistente (la aplicación es muy simple) y se puede considerar dentro de las HH del ingeniero.

	Costo Unitario [UF]	Uso [%]	Costo Final [UF]
HH Ingeniero	65	80%	52
Programador	100	100%	100
<b>Total</b>	<b>165</b>	<b>92%</b>	<b>152</b>

**Tabla 27: Inversión Inicial Proyecto. Fuente: Elaboración Propia.**

### 7.6.3. Beneficios

Los beneficios, como se discutió anteriormente, provienen del aumento de los clientes (pólizas) producto de la baja del precio, es decir, el diferencial de primas percibidas debido al aumento de los asegurados a un valor precio menor.

Finalmente, si el proyecto funciona como es debido, se espera una reducción del 11% de la prima inventario (prima dedicada a los siniestros, precio técnico), lo que se traduce en una reducción del 7% del precio final.

Se utilizan dos supuestos principales; todo se mantiene constante (esfuerzos de marketing, regulaciones, leyes, etc.) menos el precio y el parque vehicular y la composición de la prima (cuanto de ella se va a margen) se mantiene igual que al día de hoy.

Tal vez se podría agregar como beneficio el ahorro de tiempo de quien ejecuta actualmente el proceso de SOAP (de forma optimista, una reducción de un par de días de trabajo), pero al valorizar estas horas, el beneficio máximo es despreciable comparado con las primas extra percibida en el proyecto.

	Polizas_Act	Polizas_Pro	Dif
Periodo 1	590.129	779.162	189.033
Periodo 2	594.260	784.616	190.356
Periodo 3	598.420	790.108	191.689

**Tabla 28: Nuevas Pólizas Por Periodo. Fuente: Elaboración Propia.**

	Año 1	Año 2	Año 3
Ingresos [UF]	49.619	49.966	50.316

**Tabla 29: Ingresos Proyecto. Fuente: Elaboración Propia.**

#### 7.6.4. Costos

El principal costo del proyecto serán los siniestros (gastos médicos, muertes o invalideces de los involucrados) de las nuevas pólizas vendidas. Además, se asocia una comisión promedio a las pólizas, esta comisión es menor que el resto de la cartera, ya que estas pólizas intentaran ser vendidas por canal con menor comisión (como el canal web). Finalmente se contabilizan los “gastos”, concepto que calcula mensualmente control financiero e involucra todo tipo de gastos asociados a una póliza, desde su mantención física o en sistemas de datos, hasta las remuneraciones de las personas que trabajan en ella.

Para efecto de la evaluación económica, se crean tres escenarios distintos, que difieren en el número de siniestros que ocurrirán realmente, estos escenarios varían en un 5% de los siniestros esperados.

Los siniestros son calculados en base a la prima inventario que define el proyecto, esta prima cubre los siniestros futuros, por lo que si le creemos será exactamente igual a los siniestros.

Otros costos, como la mantención del computador o alguna capacitación extra, no han sido considerado al ser despreciables al compararlos con los costos producto de los siniestros esperados.

	Pesimista			Optimista			Esperado		
	Año 1	Año 2	Año 3	Año 1	Año 2	Año 3	Año 1	Año 2	Año 3
Siniestros [UF]	40.667	40.951	41.238	36.794	37.051	37.311	38.730	39.001	39.274
Comision [UF]	3.884	3.911	3.938	3.884	3.911	3.938	3.884	3.911	3.938
Gastos [UF]	3.970	3.997	4.025	3.970	3.997	4.025	3.970	3.997	4.025
<b>Total</b>	<b>48.520</b>	<b>48.860</b>	<b>49.202</b>	<b>44.647</b>	<b>44.960</b>	<b>45.274</b>	<b>46.584</b>	<b>46.910</b>	<b>47.238</b>

Tabla 30: Costos Proyecto. Fuente: Elaboración Propia.

#### 7.6.5. Flujo de Caja

Para el flujo de caja propiamente tal, se utilizó un horizonte de 3 años, principalmente debido a los posibles cambios organizacionales que dejen obsoleto el proyecto o disminuyan el interés en él.

Para la tasa de descuento, se utilizó el modelo CAPM para valorizar activos financieros:

$$E(r_j) = r_f + \beta_{jm} E(r_m - r_f) \quad (1)$$

Donde,

$E(r_j)$	es la tasa de rendimiento esperada de capital invertido en el capital j.
$\beta_{jm}$	es la beta de la industria.
$E(r_m - r_f)$	es el exceso de rentabilidad de la cartera de mercado.
$r_f$	es el rendimiento libre de riesgo.
$r_m$	es el rendimiento del mercado.

Para el caso del proyecto se utilizó una tasa libre de riesgo del 3,5%, una tasa de mercado del 10,09% y un beta de la industria (seguros generales) del 0.92. Con estos datos se obtuvo una tasa de descuento del 9.5628%.

Debido a que se tienen 3 escenarios para los costos, también existen 3 flujos de caja asociados. Para el caso del proyecto no se puede depreciar la inversión inicial.

Los indicadores económicos (en especial el ROI y el TIR) tienen valores particularmente altos, pero estos se deben principalmente a que la inversión es muy pequeña dado los beneficios potenciales, a pesar de lo extraño que parece al principio, estos valores no son tan extraños en proyectos de cambio tecnológicos como el descrito.

**Escenario más probable**

Item	Año 0	Año 1	Año 2	Año 3
(+) Ingresos		49.619	49.966	50.316
(-) Costo Siniestros		(38.730)	(39.001)	(39.274)
(-) Comision y Gastos		(7.853)	(7.908)	(7.964)
Utilidad antes de impuestos		3.035	3.056	3.078
(-) Impuesto Renta 20%		(607)	(611)	(616)
Utilidad después de impuesto		2.428	2.445	2.462
<b>(-) Inversión</b>	<b>(152)</b>			
<b>Totales</b>	<b>(152)</b>	<b>2.428</b>	<b>2.445</b>	<b>2.462</b>

**Indicadores financieros**

<b>VAN</b>	
<b>10%</b>	5.387
<b>12%</b>	5.105
<b>14%</b>	4.843
<b>16%</b>	4.600
<b>18%</b>	4.373
<b>20%</b>	4.162
<b>ROI</b>	488952%
<b>TIR</b>	1598%

Tabla 31: Flujo de Caja Proyecto (Escenario más probable). Fuente: Elaboración Propia.

**Escenario Optimista**

Item	Año 0	Año 1	Año 2	Año 3
(+) Ingresos		49.619	49.966	50.316
(-) Costo Siniestros		(36.794)	(37.051)	(37.311)
(-) Comision y Gastos		(7.853)	(7.908)	(7.964)
Utilidad antes de impuestos		4.972	5.007	5.042
(-) Impuesto Renta 20%		(994)	(1.001)	(1.008)
Utilidad después de impuesto		3.977	4.005	4.033
(-) Inversión	(152)			
<b>Totales</b>	<b>(152)</b>	<b>3.977</b>	<b>4.005</b>	<b>4.033</b>

**Indicadores financieros**

VAN	
10%	8.913
12%	8.449
14%	8.019
16%	7.618
18%	7.246
20%	6.898
ROI	800912%
TIR	2617%

**Tabla 32: Flujo de Caja Proyecto (Escenario Optimista). Fuente: Elaboración Propia.**

**Escenario pesimista**

Item	Año 0	Año 1	Año 2	Año 3
(+) Ingresos		49.619	49.966	50.316
(-) Costo Siniestros		(40.667)	(40.951)	(41.238)
(-) Comision y Gastos		(7.853)	(7.908)	(7.964)
Utilidad antes de impuestos		1.099	1.106	1.114
(-) Impuesto Renta 20%		(220)	(221)	(223)
Utilidad después de impuesto		879	885	891
(-) Inversión	(152)			
<b>Totales</b>	(152)	879	885	891

**Indicadores financieros**

VAN	
10%	1.862
12%	1.761
14%	1.668
16%	1.581
18%	1.501
20%	1.426
ROI	176993%
TIR	577%

**Tabla 33: Flujo de Caja Proyecto (Escenario Pesimista). Fuente: Elaboración Propia.**

Algunas consideraciones finales son que el proyecto es financiado por capitales propios, que no se detallan los planes de marketing y ventas debido a que escapa del alcance del proyecto mientras que los de operaciones y recursos humanos no son innecesarios, dado que la operación es sumamente simple y la gente involucrada es muy poca.

## Conclusión

En la presente tesis se trató un problema sumamente interesante y completamente vigente en la industria aseguradora Chilena, como es la mejora y automatización de procesos técnicos de un producto en específico (SOAP). Mediante la metodología de ingeniería de negocios desarrollada por el Dr. Oscar Barros, se realizó un análisis que comienza con alcances estratégicos hasta llegar a un rediseño de procesos con su correspondiente soporte tecnológico y plan de cambio.

El mercado asegurador chileno es un mercado sumamente competitivo, con fuertes actores nacionales y varios multinacionales, por lo que cualquier proyecto que logre alguna ventaja o diferenciación con respecto a la competencia tiene un enorme potencial.

Consortio Financiero es un actor muy relevante en el mercado chileno de los servicios financieros, con enormes desafíos en el futuro, por lo que la implementación correcta de un proyecto de las características del descrito en el presente informe puede provocar grandes beneficios para la empresa. Consortio es líder del mercado de Vida, pero un actor pequeño en el de generales (no-vida), esta es una de las razones de la existencia del proyecto, ya que se tiene un plan de crecimiento muy ambicioso, pero siempre manteniendo la rentabilidad esperada por el directorio.

Realizando un análisis estratégico de la empresa, se puede concluir que existe una gran oportunidad para la implementación de un proyecto que haga un uso intensivo de las TI en relación al seguro obligatorio de accidentes personales (SOAP), esto se debe principalmente a que el mercado trata este producto como un commodity, solo diferenciado por el precio y los resultados fluctuantes de las empresas demuestran que la selección de clientes no ha sido la óptima.

Por esta razón es que el principal objetivo del proyecto es el “aumento del beneficio y marketshare del SOAP mediante la automatización de procesos segmentación de clientes y pricing”, ya que al mejorar el proceso de segmentación se puede acceder a mayor información de los clientes y por ende crear ofertas diferenciadas a los mismos y de esta forma aumentar la participación de mercado en base a los clientes más rentables.

De acuerdo a la metodología de ingeniería de negocios, se generó un modelo de negocios, la justificación económica del proyecto y un diseño de los procesos requeridos por este modelo. Además, se diseñaron las aplicaciones computacionales de apoyo a los procesos y un prototipo de implementación de éstas.

La implementación del diseño propuesto tiene el potencial de aumentar los beneficios percibidos por cliente, así como crear carteras con baja siniestralidad las cuales prefieran a la empresa durante el tiempo, además en el proyecto se plantea la automatización de la gestión lo que provoca una reducción de los costos administrativos y por ende una ventaja operacional.

Los objetivos específicos del proceso provocaron varias mejoras apreciables como la creación y mantención de un Data Warehouse propio del producto SOAP, la mejora de la estimación de siniestros no reportados, la automatización de proceso de recopilación, carga y proceso de datos, así como la generación de informes periódicos de los resultados del producto.

El cálculo de los OYNR fue uno de los objetivos particulares del proyecto, particularmente debido a que el cálculo normativo dista bastante de la realidad, lo que provoca inexactitudes en las revisiones del producto así como aumenta la incertidumbre en la estimación del precio determinado a cobrar a los clientes. Por esta razón se desarrolló un modelo mixto al normativo (metodología de los triángulos), que adaptado a las condiciones del producto, produjo mejores resultados, lo que se puede apreciar en el análisis de los indicadores del back-testing realizado.

Error Porcentual Promedio (MPE)	-5,66%
Error Porcentual Absoluto (MAPE)	6,09%
Precision Estimacion Promedio	93,91%

**Tabla 7: Indicadores Backtesting OYNR. Fuente: Elaboración Propia.**

Con respecto al Clustering se determinó un número óptimo de 5 segmentos, basado en el indicador Davies-Bouldin (estabilidad de clusters), lo que se pueden apreciar a continuación:

Indicador/Nº Clusters	3	5	7	9	11
Davies-Bouldin	1,068	0,943	0,984	1,050	1,029

**Tabla 8: Indicadores Número de Clusters. Fuente: Elaboración Propia.**

Para analizar la calidad de la segmentación final se calculó el coeficiente Silhouette para el algoritmo K-means el que entrego un valor de 0,843 (siendo el valor 1 una segmentación perfecta). Esta segmentación se utilizó para calcular la prima técnica teórica de cada segmento, lo que se utiliza como medida de comparación con la tarificación original y a la vez ayuda a mejorar el pricing de los nuevos periodos.

La segmentación entrego varios resultados relevantes, como por ejemplo la identificación de las variables que explican de mejor forma el comportamiento de los clientes, lo que repercute inmediatamente en la tarificación del producto, así como permite mejorar las campañas de marketing (que regalos asociar al producto, con que empresas asociarse, en donde hacer publicidad, etc.) y elegir mejor los canales de venta (canales propios, agentes, corredores, etc.).

Por otro lado se generó la posibilidad de predecir el comportamiento de nuevos clientes y por ende la posibilidad de atraer (o retener) a los mejores clientes mediante los descuentos apropiados. Este desarrollo se basó en la implementación de redes neuronales para la clasificación de clientes entre; Malos, Riesgosos o Buenos. Para analizar el desempeño del algoritmo de clasificación se calculó la sensibilidad y especificidad del modelo, los cuales se pueden apreciar a continuación:

Data	Medida	Clase		
		Malo	Riesgoso	Bueno
Entrenamiento	Sensibilidad	75.6%	96.9%	93.4%
Entrenamiento	Especificidad	99.6%	95.8%	97.6%
Prueba	Sensibilidad	54.6%	97.5%	81.2%
Prueba	Especificidad	99.5%	96.5%	98.1%

**Tabla 14: Sensibilidad y Especificidad Red Neuronal. Fuente: Elaboración Propia.**

El sistema final se presentó como un prototipo con interfaz web, que en pocos pasos y con baja interacción del usuario final, puede procesar los nuevos datos del mes y actualizar la información utilizada en el análisis. Se materializó una disminución del 75% en el tiempo de proceso de la información del SOAP, así como en la preparación de los informes propios del producto que son enviados a diversas áreas.

La implementación final ha quedado a cargo del área de proyectos TI de Consorcio, quienes en base a los diagramas UML y al prototipo mismo crearan una versión final que sea compatible con las demás aplicaciones y servicios creados dentro de la empresa. Asimismo, esta misma área debe modificar el servicio de venta web, canal en el Consorcio ya es líder, en donde se tienen grandes esperanzas de rápidamente poder obtener buenos resultados gracias a la nueva selección de clientes.

El proyecto fue evaluado económicamente para poder establecer cuál era la factibilidad de su implementación, encontrando resultados favorables en todos los escenarios. El escenario más probable presenta un VAN superior a las 5000 UFs. La evaluación final del proyecto es difícil de realizar, ya que los resultados del producto pueden variar mucho por variables exógenas, pero al menos de forma preliminar, el proyecto ha tenido muy buena acogida, ya que las áreas de tarificación y comercialización del producto han reaccionado muy bien a la nueva información y a las nuevas herramientas que les fueron entregadas gracias al proyecto.

Dadas las características únicas del SOAP y a la estrategia de la empresa, se detectaron dos formas de obtener ventajas con respecto a la competencia y estas son; diferenciación mediante el precio, utilización de canales y socios adecuados y por otra parte mejorar la relación con el cliente (potenciar la marca), esto se puede lograr mejorando la gestión del siniestro y el proceso de post venta en general.

La primera parte fue acogida en el presente proyecto, mientras que la segunda queda propuesta, pero ya han surgido varias ideas que podrían ser muy relevantes, como por ejemplo aplicar text mining a los reclamos obtenidos mediante los diversos canales que Consorcio tiene para sus clientes, hacer social mining (sentimental mining) en las redes en que Consorcio participa (Twitter, Facebook, Youtube), para monitorear y descubrir nuevas necesidades de los clientes.

Como otra arista, se puede trabajar en la detección de fraude, del que, por experiencia internacional, se sabe de su existencia pero no ha podido ser abordado de buena forma por Consorcio debido a falta de sistemas y experiencia.

Con respecto a otros desarrollos, la implementa exitosa del proyecto sienta las base para su replicación en otros productos (y procesos) relevantes para la empresa que presentan desafíos similares, como es el caso de los seguros de salud colectivos, diversos productos de comercialización masiva o la gestión del reaseguro.

## Bibliografía

- [1] [www.consortio.cl](http://www.consortio.cl), Sitio Web de la empresa CONSORCIO SA.
- [2] Reporte Financiero 2013, Consorcio Financiero.
- [3] [www.svs.cl](http://www.svs.cl), Sitio de la Superintendencia de Valores y Seguros
- [4] Specia Jimenez, A. “Análisis jurídico de la intermediación del contrato de seguro (Capítulo 2)”. Universidad de las Américas Puebla. 2005.
- [5] [www.fundacionmapfre.org](http://www.fundacionmapfre.org), Sitio aseguradora Mapre.
- [6] Barros, O. “Ingeniería de Negocios. Diseño Integrado de Negocios, Procesos y Aplicaciones TI” (Versión 5). Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial. 2012.
- [7] Hax, A. “The Delta Model: Reinventing Your Business Strategy”. Sloan School of Management, MIT. 2010.
- [8] Porter, M., “Estrategia Competitiva”. Compañía Editorial Continental, 1995
- [9] Barros, O. (2006). La Ingeniería de Negocios y Enterprise Architecture. CEGES, Departamento de Ingeniería Industrial.
- [10] Scmuller, J. “Aprendiendo UML en 24 horas”, Prentice-Hall, 2001
- [11] Lopes, H., Et al. “A Non-parametric Method For Incurred But Not Reported Claim Reserve Estimation”. International Journal for Uncertainty Quantification (2012).
- [12] Sanchez, J. “Principios Sobre Bases de Datos Relacionales”. Stanford, California, 2004
- [13] Awad, Elias, “Systems Analysis and Design”, Second Edition, Richard D. Irwin, Inc. 1985
- [14] Adamson, C. “Star Schema The Complete Reference”, McGraw Hill, 2010
- [15] Inmon, B. “Building the Data Warehouse”. Wiley. 1992.
- [16] Kimball, R. “The Data Warehouse Toolkit”. Wiley. 1996.
- [17] Fayyad, U. Piattetsky-Shapiro, G. Smyth, P. "From Data Mining to Knowledge Discovery in Databases". 1996.

- [18] Kantardzic, M. "Data Mining: Concepts, Models, Methods, and Algorithms". John Wiley & Sons. 2003.
- [19] <http://urtechfriend-paperpresentations1.blogspot.com/p/data-mining-and-warehousing.html>
- [20] Han, J., Kamber, M. "Data Mining: Concepts And Techniques". Morgan Kaufmann. San Francisco. 2006
- [21] Alpaydin, E. "Introduction to Machine Learning". MIT Press. 2010.
- [22] Modulo "Matematicas de Seguros", Diplomado Internacional de Reaseguro. Escuela de Seguros. Chile. 2014.
- [23] Guo L, "Applying Data Mining Techniques in Property/Casualty Insurance". University of Central Florida. 2004
- [24] <http://dataconomy.com/what-is-a-data-warehouse/>
- [25] Kotter, J.P, "Leading Change: Why Transformation Efforts Fail". Harvard Business School. 1995
- [26] Curso "Innovación y Gestión de Negocios". Magister en Ingeniería de Negocios. Universidad de Chile. 2014.
- [27] Roosevelt C. Mosley. "Data Mining Applied to Insurance". Casualty Actuarial Society. 2012.
- [28] Data Mining Specialization, "Cluster Analysis in Data Mining". University of Illinois Urbana-Champaign. 2014.

## Anexos

### A. Estadísticas completas SOAP-SVS<sup>9</sup>

#### A . NUMERO DE SINIESTROS DENUNCIADOS DEL PERIODO (entre el 1 de enero y 31 de diciembre de 2012)

Sociedad	Siniestros rechazados (1)	Siniestros en revisión (2)	Siniestros aceptados (3)	Total de siniestros del período (1)+(2)+(3)
Aseguradora Magallanes	2		4.831	4.833
Bci	26		9.504	9.530
Chilena Consolidada	7		989	996
Consortio Nacional			1.757	1.757
HDI			5	5
Liberty			249	249
Mapfre	15		1.755	1.770
Mutual de Seguros			33	33
C.S.G. Penta Security	111		7.656	7.767
Renta Nacional	1	25	716	742
RSA			1.632	1.632
SURA				0
Zenit			206	206
<b>TOTAL</b>	<b>162</b>	<b>25</b>	<b>29.333</b>	<b>29.520</b>

<sup>9</sup> Todos los valores monetarios corresponden a UFs

**B. NUMERO DE SINIESTROS PAGADOS O POR PAGAR DEL PERIODO**  
 (entre el 1 de enero y 31 de diciembre de 2012)

Sociedad	Siniestros pagados (4)	Siniestros parc. pagados (5)	Siniestros por pagar (6)	Total siniestros (4)+(5)+(6)
Aseguradora Magallanes	4.333		498	4.831
Bci	3.935	5.050	519	9.504
Chilena Consolidada	439	511	39	989
Consortio Nacional	1.505	22	230	1.757
HDI	5			5
Liberty	123	116	10	249
Mapfre	837	700	218	1.755
Mutual de Seguros	23		10	33
C.S.G. Penta Security	2.815	4.453	388	7.656
Renta Nacional	701	15		716
RSA	553	962	117	1.632
SURA				0
Zenit	107	82	17	206
<b>TOTAL</b>	<b>15.376</b>	<b>11.911</b>	<b>2.046</b>	<b>29.333</b>

## C. NUMERO DE PERSONAS SINIISTRADAS DEL PERIODO

(entre el 1 de enero y 31 de diciembre de 2012)

Sociedad	Fallecidos (7)	Incapacidad permanente		Sólo gastos de hospital y otros (10)	Por siniestros en revisión (11)	Total personas siniestradas del período (7)+(8)+(9)+(10)+(11)	
		total (8)	parcial (9)				
Aseguradora Magallanes	436	41	26	503	6.799	7.805	
Bci	569	19	17	605	15.754	16.964	
Chilena Consolidada	61	6	3	70	1.156	1.296	
Consortio Nacional	62			62	1.945	2.069	
HDI				0	5	5	
Liberty	7		1	8	283	299	
Mapfre	67	5	2	74	2.469	2.617	
Mutual de Seguros	1			1	29	34	
C.S.G. Penta Security	411	23	13	447	12.360	13.254	
Renta Nacional	48	2	1	51	742	874	
RSA	91	6	2	99	2.367	2.565	
SURA				0		0	
Zenit	7	0	1	8	298	314	
				0			
<b>TOTAL</b>	<b>1.760</b>	<b>102</b>	<b>66</b>	<b>1928</b>	<b>44.207</b>	<b>33</b>	<b>48.096</b>

## D. SINIESTROS PAGADOS DIRECTOS EN EL PERIODO

(entre el 1 de enero y 31 de diciembre de 2012, montos expresados en miles de pesos de diciembre de 2012)

Sociedad	Indemnizaciones (sin gastos de hospital)			Gtos. hospital y otros (13)	Costo de liquidación (14)	Total siniestros pagados directos (12)+(13)+(14)	
	Fallecidos (12)	Inval. parcial (12)	Inval. total (12)				
Aseguradora Magallanes	1.707.563	55.324	175.555	1.938.442	2.105.347	4.043.789	
Bci	3.411.082	75.292	280.941	3.767.315	5.366.660	9.139.539	
Chilena Consolidada	462.069	8.353	40.792	511.214	487.801	999.015	
Consortio Nacional	462.224	2.294	6.752	471.270	812.927	1.284.197	
HDI				0	12.724	12.724	
Liberty	61.424	3.940		65.364	108.226	174.158	
Mapfre	385.370	17.898	100.513	503.781	952.648	1.456.429	
Mutual de Seguros	6.851			6.851	9.058	15.909	
C.S.G. Penta Security	2.592.483	97.777	325.079	3.015.339	4.358.926	7.399.627	
Renta Nacional	316.975	33.995	22.336	373.306	483.476	856.782	
RSA	568.401	11.414	126.208	706.023	837.953	1.543.976	
SURA				0		0	
Zenit	41.113	3.462	6.852	51.427	85.187	136.614	
<b>TOTAL</b>	<b>10.015.555</b>	<b>309.749</b>	<b>1.085.028</b>	<b>11.410.332</b>	<b>15.620.933</b>	<b>31.494</b>	<b>27.062.759</b>

## E. COSTO DE SINIESTROS DIRECTOS DEL PERIODO

(entre el 1 de enero y 31 de diciembre de 2012, montos expresados en miles de pesos de diciembre de 2012)

Sociedad	Siniestros directos		Reserva Ocurridos y no reportados (17)	Siniestros por pagar directos período anterior (18)	Costo siniestros directos del período (15+16+17)-(18)
	pagados (15)	por pagar (16)			
Aseguradora Magallanes	4.043.789	2.332.202	411.147	1.657.072	5.130.066
Bci	9.139.539	1.667.419	2.175.088	870.835	12.111.211
Chilena Consolidada	999.015	93.977	96.380	103.630	1.085.742
Consortio Nacional	1.284.197	214.214	480.238	109.650	1.868.999
HDI	12.724	554	3.647	4.365	12.560
Liberty	174.158	21.741	40.941	33.891	202.949
Mapfre	1.456.429	565.885	125.804	578.988	1.569.130
Mutual de Seguros	15.909	1.251	10.439		27.599
C.S.G. Penta Security	7.399.627	1.032.817	2.173.553	1.069.702	9.536.295
Renta Nacional	856.782	9.327	131.120	102.809	894.420
RSA	1.543.976	311.202	258.315	309.700	1.803.793
SURA	0				0
Zenit	136.614	32.741	47.521	4.938	211.938
<b>TOTAL</b>	<b>27.062.759</b>	<b>6.283.330</b>	<b>5.954.193</b>	<b>4.845.580</b>	<b>34.454.702</b>

## F. NUMERO DE SEGUROS CONTRATADOS

(entre el 1 de enero y 31 de diciembre de 2012)

Sociedad	Automóviles	Camionetas	Motocicletas	Otros	Camiones	Buses	Taxis	Total
Aseguradora Magallanes	466.017	142.290	6.584	48.600	3.624	31.232	485	698.832
Bci	746.398	338.655	50.406	59.359	65.309	29.227	26.589	1.315.943
Chilena Consolidada	136.781	49.574	8.875	3.756	52	3	3	199.044
Consortio Nacional	364.107	85.136	7.311	4.784	2.985	1.770	1.408	467.501
HDI	451	59		174				684
Liberty	6.657	707	7		508	3	7	7.889
Mapfre	223.791	48.260	6.496	8.927	9.409	5.998	921	303.802
Mutual de Seguros	3.729	773		23				4.525
C.S.G. Penta Security	330.184	270.850	28.554	22.250	110.864	16.651	62.619	841.972
Renta Nacional	31.074	28.087	14	1.911	170	9.310	828	71.394
RSA	160.161	45.343	11.859	13.490	18.602	10.532	15.514	275.501
SURA								0
Zenit	24.947	18.814	559	143				44.463
<b>TOTAL</b>	<b>2.494.297</b>	<b>1.028.548</b>	<b>120.665</b>	<b>163.417</b>	<b>211.523</b>	<b>104.726</b>	<b>108.374</b>	<b>4.231.550</b>

## G. PRIMA TOTAL POR TIPO DE VEHICULO

(entre el 1 de enero y 31 de diciembre de 2012, montos expresados en miles de pesos de diciembre de 2012)

Sociedad	Automóviles	Camionetas	Camiones	Buses	Motocicletas	Taxis	Otros	Total
Aseguradora Magallanes	5.521.060	1.879.483	81.397	402.145	252.603	8.771	931.253	9.076.712
Bci	6.810.612	3.759.738	1.287.655	1.306.746	1.633.699	516.337	469.288	15.784.075
Chilena Consolidada	1.273.547	595.818	1.046	30	303.035	35	46.906	2.220.417
Consortio Nacional	2.977.293	942.287	63.786	31.281	262.162	29.461	28.381	4.334.651
HDI	3.311	616					1.035	4.962
Liberty	50.261	6.219	10.209	61	376	133		67.259
Mapfre	1.784.425	525.205	188.339	610.804	220.609	18.370	46.076	3.393.828
Mutual de Seguros	36.122	9.995					112	46.229
C.S.G. Penta Security	3.188.927	3.064.941	1.810.361	1.748.036	979.646	1.206.626	258.886	12.257.423
Renta Nacional	226.717	262.060	1.911	398.883	371	7.903	23.161	921.006
RSA	1.296.011	482.430	356.140	871.874	357.570	284.884	76.656	3.725.565
SURA								0
Zenit	184.483	171.707			19.333		595	376.118
<b>TOTAL</b>	<b>23.352.769</b>	<b>11.700.499</b>	<b>3.800.844</b>	<b>5.369.860</b>	<b>4.029.404</b>	<b>2.072.520</b>	<b>1.882.349</b>	<b>52.208.245</b>

## H. PRIMA PROMEDIO POR TIPO DE VEHICULO

(entre el 1 de enero y 31 de diciembre de 2012, montos expresados en miles de pesos de diciembre de 2012)

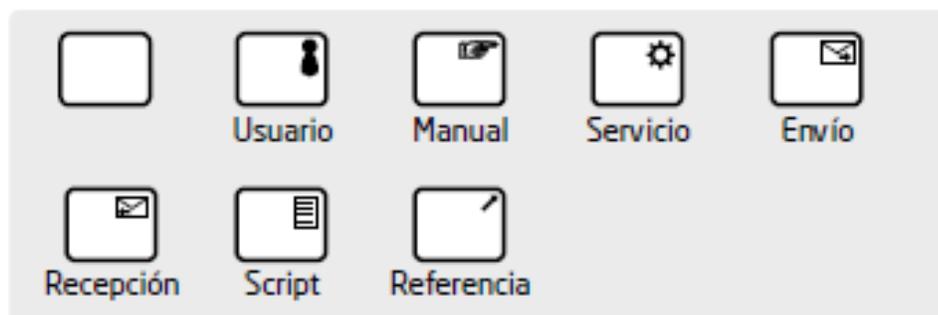
Sociedad	Automóviles	Camionetas	Camiones	Buses	Motocicletas	Taxis	Otros	Promedio
Aseguradora Magallanes	11.847	13.209	22.461	12.876	38.366	18.085	19.162	12.988
Bci	9.125	11.102	19.716	44.710	32.411	19.419	7.906	11.994
Chilena Consolidada	9.311	12.019	20.115	10.000	34.145	11.667	12.488	11.155
Consortio Nacional	8.177	11.068	21.369	17.673	35.859	20.924	5.932	9.272
HDI	7.341	10.441					5.948	7.254
Liberty	7.550	8.796	20.096	20.333	53.714	19.000		8.526
Mapfre	7.974	10.883	20.017	101.835	33.961	19.946	5.161	11.171
Mutual de Seguros	9.687	12.930					4.870	10.216
C.S.G. Penta Security	9.658	11.316	16.330	104.981	34.309	19.269	11.635	14.558
Renta Nacional	7.296	9.330	11.241	42.845	26.500	9.545	12.120	12.900
RSA	8.092	10.640	19.145	82.783	30.152	18.363	5.682	13.523
SURA								-
Zenit	7.395	9.127			34.585		4.161	8.459
<b>PROMEDIO</b>	<b>9.362</b>	<b>11.376</b>	<b>17.969</b>	<b>51.275</b>	<b>33.393</b>	<b>19.124</b>	<b>11.519</b>	<b>12.338</b>

## B. Notación BPMN

- **Actividades**

### Tarea

Son actividades simples o atómicas. No es definida a un nivel más detallado. Existen diferentes tipos:



### Subproceso

Es una actividad compuesta que incluye un conjunto interno lógico de actividades (proceso) y que puede ser analizado en más detalle.

 **Subproceso embebido**  
Depende del proceso padre.  
No puede contener pools ni lanes.

 **Subproceso reusable**  
Es un proceso definido como un diagrama de procesos independiente y que no depende del proceso padre.

- **Compuertas**

**Compuerta Exclusiva basada en datos**

**Divergencia:** Ocurre cuando en un punto del flujo basado en los datos del proceso se escoge un solo camino de varios disponibles.

**Convergencia:** Como punto de convergencia, es utilizada para confluir caminos excluyentes.

**Compuerta Exclusiva basada en eventos**

La compuerta exclusiva basada en eventos representa un punto del proceso donde se escoge un camino de varios disponibles, pero la decisión no se basa en datos del proceso sino en eventos.

**Compuerta Paralela**

**Divergencia:** Se utiliza cuando varias actividades pueden realizarse concurrentemente o en paralelo.

**Convergencia:** Permite sincronizar varios caminos paralelos en uno solo. El flujo continúa cuando todos los flujos de secuencia de entrada hayan llegado a la figura.

**Compuerta Inclusiva**

**Divergencia:** Se utiliza cuando en un punto se activan uno o más caminos de varios caminos disponibles, basado en los datos del proceso.

**Convergencia:** Se utiliza para sincronizar caminos activados previamente por una compuerta inclusiva usada como punto de divergencia.

**Compuerta Compleja**

**Divergencia:** Es utilizada para controlar puntos de decisión complejos.

**Convergencia:** permite continuar al siguiente punto del proceso cuando una condición de negocio se cumple.

- **Eventos de Inicio**

**Evento de Inicio sin especificar**

No se especifica ningún comportamiento en particular para iniciar el proceso.

**Evento de Inicio de Mensaje**

Un proceso inicia cuando un mensaje es recibido.

**Evento de Inicio de Temporización**

Indica que un proceso inicia cada ciclo de tiempo o en una fecha específica.

**Evento de Inicio de Condición**

Un proceso inicia cuando una condición de negocio se cumple.

**Evento de Inicio de Señal**

El proceso inicia cuando se captura una señal lanzada desde otro proceso. Tenga en cuenta que una señal no es un mensaje, un mensaje tiene claramente definido un destinatario, la señal no.

**Evento de Inicio Múltiple**

Indica que existen muchas formas de iniciar el proceso y que al cumplirse una de ellas se iniciará el proceso.

- **Eventos Intermedios**


**Evento Intermedio sin especificar**

Indica algo que ocurre o puede ocurrir dentro del proceso, sólo se pueden utilizar dentro de la secuencia del flujo.


**Evento Intermedio de Mensaje**

Indica que un mensaje puede ser enviado o recibido.



Si el evento de mensaje es de recepción, indica que el proceso no continúa hasta que el mensaje sea recibido.

Puede utilizarse dentro del flujo de secuencia o adjunto a los límites de una actividad para indicar un flujo de excepción.


**Evento Intermedio de Temporización**

Indica una espera dentro del proceso. Este tipo de evento puede utilizarse dentro del flujo de secuencia indicando una espera entre las actividades o adjunto a los límites de una actividad indicando un flujo de excepción.


**Evento Intermedio de Condición**

Se utiliza para esperar que una condición de negocio se cumpla. Se puede utilizar dentro del flujo de secuencia indicando que se espera a que la condición de negocio se cumpla o adjunto a los límites de una actividad indicando un flujo de excepción que se activará cuando la condición se cumpla.


**Evento Intermedio de Señal**

Se utiliza para enviar o recibir señales. Se puede utilizar dentro del flujo de secuencia para enviar o recibir señales o adjunto a los límites de una actividad indicando un flujo de excepción que se activará cuando la señal sea capturada.


**Evento Intermedio Múltiple**

Indica que puede ser activado por muchas causas.


**Evento Intermedio de Cancelación**

Este tipo de evento intermedio es usado en subprocesos Transaccionales. Se diagrama a los límites del Subproceso transaccional indicando un flujo alternativo que se realizaría cuando el subproceso transaccional es cancelado. Se diagrama a los límites del subproceso.


**Evento Intermedio de Error**

Esta figura es usada para capturar errores. Se diagrama a los límites de una actividad.


**Evento Intermedio de Compensación**

Permite manejar compensaciones, cuando se utiliza dentro del flujo de secuencia de un proceso indica que una compensación es necesaria, es decir se lanza una compensación. Cuando se utiliza adjunto a los límites de una actividad indica que esa actividad será compensada cuando se active el evento.


**Evento Intermedio de Enlace**

Este evento permite conectar dos secciones del proceso.



- **Eventos de Finalizacion**

**Evento de Fin sin especificar**

Indica que un camino del flujo llega al fin.

**Evento de Fin de Mensaje**

Permite enviar un mensaje al finalizar el flujo.

**Evento de Fin de Señal**

Permite enviar una señal al finalizar el flujo.

**Evento de Fin Múltiple**

Indica que varios resultados pueden darse al finalizar un flujo.

**Evento de Fin de Cancelación**

Permite enviar una excepción de cancelación al finalizar el flujo. Sólo se utiliza en subprocesos transaccionales.

**Evento de Fin de Error**

Permite enviar una excepción de error al finalizar el flujo.

**Evento de Fin de Compensación**

Este tipo de fin indica que es necesaria una compensación al finalizar el flujo.

**Evento de Fin de Terminal**

Indica que el proceso es terminado, es decir cuando algún camino del flujo llega a este fin el proceso termina completamente, sin importar que existan más caminos del flujo pendientes.

## C. Venta Web SOAP

Compra tu SOAP
 Información del Producto
 Imprime tu Póliza
 Modifica tus Datos
 Consulta de Vigencia
 ¿Problemas?

### Compra tu SOAP en línea

Paso 1 : Datos de tu SOAP
Paso 2 : Carro de Compra
Paso 3 : Pago
Paso 4 : Comprobante

#### Datos Personales

*Todos los datos son obligatorios*

RUT Pagador  Ej: 11111111-1

#### Datos del Vehículo

Ingresando la Patente se cargarán todos los demás datos, de no ser así complétalos

Patente  -  Ej: AA11111-1, AAAA11-1

**A** Marca  Ej: Toyota, Subaru

Año  Ej: 1999

**B** Tipo Vehículo

**C** Modelo

Número Motor

#### Dónde encontrar la información



#### Datos del Propietario del Vehículo

*Todos los datos son obligatorios*

RUT  Ej: 11111111-1      Nombres

Apellido Paterno       Apellido Materno

Quiero agregar otro vehículo
Continuar

**Opcional: Despacho Póliza a Domicilio**

Igualmente solicito el despacho de la póliza a domicilio con un costo adicional por póliza de \$3.000 sobre el precio de compra antes señalado. El despacho se realizará en un plazo de tres días hábiles para Santiago y cuatro días hábiles para regiones.