

TABLA DE CONTENIDO

1. ANTECEDENTES GENERALES	1
1.1 Industria Corredora Seguros Grandes Tiendas	1
1.2 Seguros Falabella	1
2. DESCRIPCIÓN Y JUSTIFICACIÓN DEL PROYECTO	2
2.1 IDENTIFICACIÓN DEL PROBLEMA Y OPORTUNIDAD	2
3. OBJETIVOS.....	4
3.1 Objetivo general	4
3.2 Objetivos específicos	4
4. ALCANCES	4
5. RESULTADOS ESPERADOS	5
6. MARCO TEÓRICO	6
6.1 Proceso de extracción y transformación de datos	6
6.2 Modelos de propensión	6
6.2.1 Modelo Logit Binario	7
6.2.2 Modelo Árbol de Decisión	8
6.3 Next Best Offer	8
6.4 Métricas de comparación	8
7. METODOLOGÍA	10
7.1 Descripción de fuentes de datos	10
7.2 Tratamiento de datos	10
7.3 Desarrollo de modelos de propensión	10
7.4 Evaluación.....	11
7.4.1 Desempeño modelos de propensión.....	11
7.5 Propuestas de acción	11
7.6 Propuesta de diseño experimental.....	11
8. DESARROLLO.....	11
8.1 Descripción de los datos	11
8.1.1 Tenencia de seguros	11
8.1.2 Gastos tarjeta CMR.....	12
8.1.3 Datos demográficos	12
8.2 Tratamiento de datos	13
8.2.1 Selección.....	13

8.2.2	Pre-procesamiento	14
8.2.3	Transformación.....	14
8.2.4	Descripción de datos	15
8.3	Desarrollo de modelos de propensión	22
8.3.1	Modelo Logit Binario	22
8.3.2	Modelo Árbol de Decisión.....	27
8.4	Evaluación de modelos.....	31
8.4.1	Desempeño modelos de propensión.....	31
8.5	Propuestas de acción	37
8.6	Propuesta de diseño experimental.....	38
8.7	Evaluación económica	40
9.	CONCLUSIONES	42
9.1	Conclusiones del proyecto.....	42
9.2	Limitaciones y trabajos futuros	43
10.	BIBLIOGRAFÍA	45
11.	ANEXOS	47
11.1	Anexo A.....	47
11.2	Anexo B.....	47
11.3	Anexo C	49
11.4	Anexo D.....	51
11.5	Anexo E.....	75
11.6	Anexo F	75

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Aumento porcentual de primas intermediadas por compañía	2
Ilustración 2: Evolución fuga y cross-selling seguro Automotriz.....	3
Ilustración 3: Esquema proceso ETL.....	6
Ilustración 4: Distribución base entrenamiento y testeo	7
Ilustración 5: Gráfico de método de creación de Matriz de Confusión	9
Ilustración 6: Distribución de tenencia según ramo	15
Ilustración 7: Distribución de clientes por edades	16
Ilustración 8: Porcentaje de contrataciones de seguros de vida según edad clientes con CMR.....	17
Ilustración 9: Porcentaje de contrataciones de seguros de vida según edad clientes no CMR.....	17
Ilustración 10: Distribución de clientes por género	18
Ilustración 11: Distribución de clientes por GSE.....	18
Ilustración 12: Distribución de clientes por Estado Civil	19
Ilustración 13: Distribución de clientes por habitación	19
Ilustración 14: Distribución tenencia vehicular.....	20
Ilustración 15: Contratación de seguros automotriz según cantidad de vehículos para clientes con CMR.....	20
Ilustración 16: Contratación de seguros automotriz según cantidad de vehículos para clientes con CMR.....	21
Ilustración 17: Distribución de clientes según monto de gastos por rubros	22
Ilustración 18: Extracto perfil más propenso modelo auto con CMR.....	23
Ilustración 19: Extracto perfil más propenso modelo vida con CMR	24
Ilustración 20: Extracto perfil más propenso modelo vida con bonificación con CMR ...	25
Ilustración 21: Extracto perfil más propenso modelo salud con CMR.....	25
Ilustración 22: Extracto perfil más propenso modelo hogar con CMR	26
Ilustración 23: Extracto perfil más propenso modelo transaccional con CMR.....	27
Ilustración 24: Gráfico Árbol de Decisión modelo auto con CMR.....	28
Ilustración 25: Gráfico Árbol de Decisión modelo vida con CMR	28
Ilustración 26: Gráfico Árbol de Decisión modelo vida con bonificación con CMR.....	29
Ilustración 27: Gráfico Árbol de Decisión modelo salud con CMR.....	29
Ilustración 28: Gráfico Árbol de Decisión modelo hogar con CMR.....	30
Ilustración 29: Gráfico Árbol de Decisión modelo transaccionales con CMR.....	30
Ilustración 30: Probabilidad media por producto modelo Logit Binario y Árbol de Decisión sin CMR.....	31
Ilustración 31: Desempeño modelos según matrices de confusión sin CMR	31
Ilustración 32: Diferencia en métricas de desempeño Logit – Arbol sin CMR	32
Ilustración 33: Curva de ganancia modelo auto sin CMR	33
Ilustración 34: Desempeño mejor quintil sin CMR modelo Logit y Arbol	33
Ilustración 35: Probabilidad media por producto modelo Logit Binario y Árbol de Decisión con CMR.....	34
Ilustración 36: Desempeño modelos según matrices de confusión con CMR.....	34
Ilustración 37: Diferencia métricas de desempeño modelo Logit - Árbol con CMR	34
Ilustración 38: Curva de ganancia modelo auto con CMR.....	35
Ilustración 39: Desempeño mejor quintil con CMR modelo Logit y Arbol	35
Ilustración 40: Ejemplo máximo decil	38
Ilustración 41: Descripción experimento máximo decil.....	39

Ilustración 42: Descripción experimento máximo beneficio esperado	39
Ilustración 43: Ventas simuladas según modelo propuesto.....	40