

TABLA DE CONTENIDO

1. ANTECEDENTES GENERALES.....	1
1.1 Industria de servicios financieros.....	1
1.2 Descripción de la Empresa.....	4
1.3 Desempeño Organizacional	5
2. DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN	9
2.1 Información de los Datos	9
2.2 Indicación del Problema, Posibles Causas y Efectos.....	12
2.3 Propuesta de Valor a la Empresa	13
3. OBJETIVOS.....	16
3.1 Objetivo General.....	16
3.2 Objetivos Específicos	16
4. ALCANCES.....	16
5. MARCO CONCEPTUAL	17
5.1 Minería de Textos	17
5.1.1 Definiciones Básicas.....	17
5.1.2 Algoritmos de Informtion Retrieval (IR).....	18
5.1.3 Análisis de Sentimientos	20
5.1.4 Distancias	22
5.2 Modelos para la Construcción de Fuga.....	22
5.2.1 Estimación de un Modelo de Propensión	22
5.2.2 Elección del Algoritmo de Matching para encontrar los Clientes Homólogos	24
5.3 Modelos basados en Árboles	25
5.3.1 Information Gain	25
5.3.2 Arboles de Decisión	25
5.3.3 Random Forest.....	26
5.4 Otros Modelos.....	26
5.4.1 Estudio de Recency, Frecuency y Monetary Value (RFM).....	26
6. METODOLOGÍA	28
6.1 Entendimiento del negocio	28
6.2 Entendimiento de la data.....	29
6.3 Preparación de la data	29
6.4 Modelado	29
6.5 Evaluación	30

6.6	Ejecución	30
7.	RESULTADOS OBTENIDOS	31
7.1	Definición de Fuga.....	31
7.1.1	Variación del indicador dentro de un mismo periodo de tiempo	32
7.1.2	Variación del indicador con respecto al periodo de tiempo anterior	33
7.1.3	Análisis de reactivación de clientes.....	34
7.1.4	Definición de Fuga	35
7.2	Minería de texto	35
7.2.1	Pre-Procesamiento de los datos	35
7.2.2	Análisis de sensibilidad de Pitman-Yor	37
7.2.3	Resultados minería de texto.....	38
7.2.4	Análisis de sentimientos	39
7.3	Homologación de clientes.....	41
7.3.1	Modelo de propensión a tópico	41
7.3.2	Resultados del emparejamiento	41
7.4	Modelo de fuga	43
7.4.1	Unificación de la base de datos	43
7.4.2	Eliminación de Outliers	43
7.4.3	Balanceo de la base de datos	44
7.4.4	Selección de variables	45
7.4.5	Modelo de fuga tradicional.....	46
7.4.6	Modelo de fuga con variables de minería de texto.....	48
7.4.7	Modelo de fuga con clientes homologados	52
7.4.8	Selección de modelo.....	54
7.4.9	Análisis de Sensibilidad	55
7.5	Trabajos futuros	56
7.6	Líneas de acción a proponer	56
8.	CONCLUSIONES.....	58
9.	BIBLIOGRAFÍA.....	60
10.	ANEXOS	62
10.1	Anexo I: Árbol de evolución R/F	62
10.2	Anexo II: Árbol de evolución Recency	62
10.3	Anexo III : Árbol de evolución Frecuencia	63
10.4	Anexo IV: Resultados Pitman-Yor	63
10.5	Anexo V: Reglas comunes Modelo de fuga tradicional	63

10.6	Anexo VI: Reglas comunes Modelo de fuga con variables de minería de texto considerando exclusivamente a clientes con comentarios.....	64
10.7	Anexo VII: Modelo de minería de texto considerando el total de los clientes	65
10.8	Anexo VIII: Reglas comunes Modelo de fuga con variables de minería de texto generadas con logit multivariado de homologación de clientes	66