

Tabla de Contenido

1. Introducción	1
2. Identificación del problema y Objetivos	3
2.1. Sector industrial	3
2.2. Estado de la satisfacción en la empresa	6
2.3. Objetivo general	9
2.4. Objetivos específicos	9
2.5. Alcances	10
3. Marco conceptual	12
3.1. Customer relationship management	12
3.2. Machine learning	12
3.2.1. Regresión logística	13
3.2.2. Árboles de decisión	14
3.2.3. Distributed random forest	14
3.2.4. Gradient boosting machine y Extreme gradient boosting	15
3.2.5. Criterio de información de Akaike	15
3.3. Métricas de evaluación de modelos	16
3.4. Teoría de encuestas	17
3.4.1. Sesgo en las encuestas	18
3.4.2. Diseño de pesos para encuestas	19
4. Metodología	21
4.1. Metodología CRISP-DM	21
5. Desarrollo	24
5.1. Revisión de las encuestas	24
5.2. Selección y elaboración de atributos	29
5.2.1. Visitas a tiendas	30
5.2.2. Uso de la aplicación	33
5.2.3. Base resumen de interacciones	37
5.3. Elección de ventana de tiempo y criterio de satisfacción	41
5.4. Modelos de satisfacción de caja negra	48
5.5. Aplicación del modelo	56
5.6. Replicabilidad en prepago	63
6. Conclusiones	68
6.1. Recomendaciones futuras	70

Bibliografía	72
Anexo A. Variables utilizadas	74
Anexo B. Interacciones	77