

Tabla de Contenido

1. Planteamiento del problema	1
1.1. Antecedentes generales	1
1.1.1. Características de la organización	1
1.1.2. Actores relevantes	4
1.1.3. Desempeño organizacional	5
1.2. Problema y justificación	6
1.2.1. Área donde se desarrolla el trabajo	6
1.2.2. Problema identificado y su relevancia	7
1.2.3. Hipótesis y alternativas de solución	9
1.2.4. Justificación del problema	11
1.2.5. Valor generado a través de la solución	12
2. Objetivos	14
2.1. Objetivo general	14
2.2. Objetivos específicos	14
2.3. Alcances	14
3. Marco conceptual	16
3.1. Machine Learning	16
3.2. Entrenamiento y validación de modelos de Machine Learning	16
3.3. Técnicas para manejo de datos desbalanceados	18
3.3.1. Under-Sample	18
3.3.2. Over-Sample	18
3.3.3. SMOTE	19
3.4. Árboles de decisión (CART)	19
3.5. Regresión Logística	21
3.6. Métricas de desempeño de los algoritmos	22
3.6.1. Matriz de confusión	22
3.6.2. Curva ROC y AUC	24

4. Metodología	25
5. Desarrollo metodológico	27
5.1. Fase de comprensión del negocio	27
5.1.1. Entendiendo el problema u oportunidad	27
5.1.2. Los datos	29
5.2. Fase de recolección y comprensión de los datos	30
5.2.1. Base de datos	30
5.2.2. Atributos de la base	30
5.3. Fase de preparación de los datos	33
5.3.1. Limpieza de tablas	33
5.3.2. Consolidación de base de datos	34
5.3.3. Creación de atributos	36
5.3.4. Variable dependiente	41
5.3.5. Análisis descriptivo de los datos	42
5.4. Fase de modelado	47
5.4.1. Machine Learning para determinar las variables más importantes en la reactivación de los clouders	47
5.4.2. Machine Learning para conocer la propensión de reactivación de los clouders	49
5.5. Fase de evaluación	53
5.6. Fase de despliegue	55
5.6.1. Productivización del modelo	55
5.6.2. Impacto económico	56
5.6.3. Riesgos del proyecto	58
6. Conclusiones	59
6.1. Cumplimiento de objetivos	59
6.2. Hipótesis planteadas	60
6.3. Impacto del trabajo realizado	61
6.4. Cambios en la metodología	62
Bibliografía	64
Anexos	65
Anexo A.	65
Anexo B.	66

Anexo C.	69
C.1. Escenario optimista	69
C.2. Escenario medio	71
C.3. Escenario pesimista	73