

Tabla de Contenido

1. Introducción	1
2. Antecedentes generales	2
2.1. Caracterización de la empresa	2
2.2. Información del área de la empresa	3
2.3. Industria bancaria	3
3. Descripción del proyecto y justificación	7
4. Objetivos y alcance	10
4.1. Objetivo General	10
4.2. Objetivos específicos	10
4.3. Alcances	10
5. Marco conceptual	12
5.1. CRISP-DM	12
5.2. Minería de textos	13
5.2.1. Enfoques	14
5.3. Algoritmos de Information Retrieval (IR)	14
5.4. Árboles de decisión	15
5.5. Gradient tree boosting	15
5.6. XGBoost	17
5.7. Métricas de evaluación	17
5.8. NPS, Impacto y Criticidad	18
5.9. Imputación múltiple	19
6. Metodología	21
6.1. Entender el negocio	21
6.2. Entender los datos	21
6.3. Preparación de los datos	22
6.4. Modelamiento	23
6.5. Evaluación	23
6.6. Despliegue	23
7. Desarrollo metodológico	24
7.1. Entender los datos	24
7.2. Preparación de los datos	29
7.3. Modelamiento	33

7.3.1. Modelo de Tópicos	34
7.3.2. Modelo de Caracterización de Lealtad	37
7.4. Evaluación	39
7.5. Despliegue de resultados	41
8. Análisis de resultados	42
8.1. Análisis de tópicos	42
8.2. Modelo de caracterización	45
9. Conclusiones	57
9.1. Conclusiones generales	57
9.2. Recomendaciones para Ipsos	59
9.3. Recomendaciones bancarias para la muestra analizada	60
9.4. Trabajos futuros	61
Bibliografía	63