

TABLA DE CONTENIDO

1. Introducción	1
1.1 Exposición del tema	1
1.1.1 Definiciones importantes	2
1.2 Justificación	4
1.3 Objetivos	8
1.3.1 Objetivo general	8
1.3.2 Objetivos específicos	8
1.4 Alcances y limitaciones del estudio	8
2. Marco teórico	9
3. Metodología	10
3.1 Tipos de modelos	14
3.2 Comparación de resultados	15
3.2.1 Matriz de confusión	15
3.2.2 Estadístico C o AUC	16
3.3 Tecnología	17
4. Recopilación y entendimiento de los datos	18
4.1 Datos iniciales	18
4.2 Estudio de variables existentes	18
4.3 Construcción de variable de objetivo	19
5. Preparación, transformación y selección de datos	20
5.1 Limpieza de datos	20
5.2 Análisis de los datos	22
5.3 Selección de variables	23
5.4 Transformación de variables	23
6. Aplicación del modelo	25
6.1 Análisis de variables	25
6.1.1 Primera iteración	26
6.1.2 Segunda iteración	30
7. Evaluación de resultados	35
7.1 Escalamiento a score	35

7.2	Generación del scorecard	41
7.2.1	Validación del scorecard	41
7.2.2	Generación de la herramienta de admisión	42
7.3	Cálculo de los beneficios	44
7.3.1	Determinación de puntaje de corte	44
8.	Conclusiones.....	52
8.1	Conclusiones generales.....	52
8.2	Conclusiones del trabajo realizado.....	53
9.	Recomendaciones	54
10.	Bibliografía	56
11.	Anexos	57
11.1	Anexo 1: Campos seleccionados y tratamiento de outliers	57
11.2	Anexo 2: NAAO.....	59
11.3	Anexo 3: Resumen Análisis Univariado	62
11.4	Anexo 4: Filtro de variables.....	63
11.5	Anexo 5: Matriz de covarianza	65
11.6	Anexo 6: Resultados Logit	68
11.7	Anexo 7: Distribución categorización variables continuas	69
11.8	Anexo 8: Tramos por punto de corte.....	70

ÍNDICE DE TABLAS

Tabla 1: Selección de tablas consolidado Tanner S.A (Fuente: elaboración propia en base a consolidado de Tanner S.A.)	18
Tabla 2: Porcentaje de Outliers y Extremos	22
Tabla 3: Variables seleccionadas. (Fuente: elaboración propia).....	23
Tabla 4: Betas calculados para primer modelo	27
Tabla 5: Betas calculados para cada variable y categoría	33
Tabla 6: Extremos ln(Odds) originales.....	36
Tabla 7: Nuevos extremos ln(odds)	36
Tabla 8: Puntaje calculado para cada variable de acuerdo a su categoría correspondiente	38
Tabla 9: Tabla de puntos de corte por rango	45
Tabla 10: Tabla de puntos de corte	47
Tabla 11: Promedio variables operación, ingresos y costos calculados	48
Tabla 12: Utilidades promedio por rango de Score	49
Tabla 13: Scorecard final	50

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Diagrama de vínculos entre actores. (Fuente: elaboración propia).....	4
Ilustración 2: Proceso Knowledge Discovery in Databases (KDD). (Fuente: Elaboración propia en base a Fayyad, 1996)	11
Ilustración 3: Ejemplos de visualización de curva ROC.....	17
Ilustración 4: Flujo del proceso de Mora con sus posibles escenarios.....	20
Ilustración 5: Resumen de las distribuciones de las variables categorizadas	24
Ilustración 6: Prototipo funcional scorecard	43
Ilustración 7: Prototipo funcional scorecard	43
Ilustración 8: Prototipo funcional scorecard.....	44