

Índice de Contenidos

| | |
|---|----|
| 1. Introducción | 1 |
| 1.1 Problemática | 2 |
| 1.2 Objetivos | 3 |
| 1.2.1 Objetivo General | 3 |
| 1.2.2 Objetivos Específicos | 4 |
| 1.3 Alcance | 4 |
| 1.4 Metodología de la Tesis | 5 |
| 1.5 Resultados Esperados | 6 |
| 1.6 Estructura de la Tesis | 6 |
| 2. Proceso <i>KDD</i> | 8 |
| 2.1 Selección de Datos | 9 |
| 2.2 Preprocesamiento de Datos | 9 |
| 2.3 Transformación de Datos | 10 |
| 2.4 Minería de Datos | 12 |
| 2.5 Interpretación y Evaluación | 13 |
| 3. <i>Credit Scoring</i> | 14 |
| 3.1 Regresión Logística | 18 |
| 3.2 Bondad de Ajuste | 21 |
| 3.2.1 Curva ROC y AUC | 22 |
| 3.3 Validación Cruzada y Punto de Corte | 23 |
| 4. Metodología de Detección del Cambio (MDC) | 25 |
| 4.1 Estabilidad de Parámetros | 27 |
| 4.2 Capacidad Discriminante de las Variables | 29 |
| 4.2.1 <i>Test de Kolmogorov – Smirnov (K-S)</i> para dos muestras | 31 |
| 4.2.2 <i>Test Chi-Cuadrado de Pearson</i> | 32 |
| 4.3 Distribución de las Variables | 34 |
| 4.3.1 Entropía | 34 |
| 4.3.2 Divergencia de Kullback - Leibler | 36 |
| 4.3.3 Experiencias Previas | 37 |
| 4.4 Capacidad Discriminante del Modelo en su Conjunto | 38 |
| 5. Implementación de MDC | 41 |
| 5.1 Evaluación del Cliente | 42 |

| | | |
|-----------|--|-----------|
| 5.2 | Decidir Crédito..... | 43 |
| 5.3 | Mantener el Modelo de Comportamiento de Clientes..... | 44 |
| 5.3.1 | Seleccionar Set de Datos para Evaluación..... | 45 |
| 5.3.2 | Evaluar Desempeño del Modelo Original..... | 45 |
| 5.3.3 | Seleccionar Criterios de Re-calibración..... | 52 |
| 5.3.4 | Comparar Desempeño de Modelos..... | 53 |
| 5.3.5 | Reemplazar Modelo Original..... | 54 |
| 6. | Resultados Experimentales..... | 55 |
| 6.1 | Descripción de la Base de Datos..... | 55 |
| 6.2 | Medida de las Variables del Modelo..... | 56 |
| 6.2.1 | Variable Dependiente..... | 56 |
| 6.2.2 | Variables Independientes..... | 57 |
| 6.2.3 | Transformación de Variables..... | 57 |
| 6.3 | Modelo de <i>Credit Scoring</i> Original..... | 58 |
| 6.3.1 | Bondad de Ajuste y Validación del Modelo..... | 60 |
| 6.4 | Metodología de Detección del Cambio (MDC)..... | 63 |
| 6.4.1 | Análisis del Modelo Original..... | 63 |
| 6.4.2 | Alternativas de Punto de Corte..... | 69 |
| 6.4.3 | Seleccionar Criterios de Re-calibración..... | 71 |
| 7. | Conclusiones..... | 72 |
| 7.1 | Conclusiones Generales..... | 72 |
| 7.2 | Principales Resultados..... | 74 |
| 7.3 | Desarrollo de Objetivos Planteados..... | 76 |
| 7.4 | Trabajos Futuros..... | 77 |
| | Bibliografía..... | 78 |
| | Anexos..... | 82 |
| A. | Salidas SPSS..... | 82 |
| A.1 | Modelo Original (2000 – 2004)..... | 82 |
| B. | Cálculo Divergencia de <i>Kullback – Leibler</i> | 87 |
| C. | Acuerdo de Capital de Basilea II..... | 96 |
| C.1 | Basilea II y Los Tres Pilares..... | 96 |
| C.2 | Basilea II y su Implementación en Chile..... | 100 |
| D. | Descripción de Variables..... | 102 |
| D.1 | Variable Dependiente..... | 102 |
| D.2 | Variables Independientes..... | 102 |

Índice de Figuras

| | |
|---|----|
| Figura 1. Etapas del Proceso <i>KDD</i> | 8 |
| Figura 2. Curva ROC para 2 Modelos..... | 22 |
| Figura 3. Proceso de Evaluación del Cliente | 42 |
| Figura 4. Proceso de Decidir Crédito | 43 |
| Figura 5. Etapas de la MDC | 45 |
| Figura 6. Curva ROC..... | 60 |

Índice de Tablas

| | |
|--|----|
| Tabla 1. Tipos de Errores y Decisión en las Pruebas de Hipótesis | 30 |
| Tabla 2. Matriz de Frecuencias Observadas | 33 |
| Tabla 3. Matriz de Frecuencias Esperadas..... | 33 |
| Tabla 4. Matriz de Confusión para el Problema de Clasificación Binario | 39 |
| Tabla 5. Descripción Variables Independientes..... | 57 |
| Tabla 6. Porcentaje de Casos para <i>Train</i> y <i>Test</i> del Modelo Original | 59 |
| Tabla 7. Coeficientes de la Regresión Logística para el Modelo Original | 59 |
| Tabla 8. Matriz de Confusión para el Modelo Original | 61 |
| Tabla 9. Punto de Corte para Clasificación de Clientes Modelo Original | 62 |
| Tabla 10. Resultados Estabilidad de Parámetros | 63 |
| Tabla 11. Resultados Análisis Univariado..... | 65 |
| Tabla 12. Resultados Divergencia de <i>Kullback – Leibler</i> | 66 |
| Tabla 13. Resumen Detección de Cambios en Variables | 67 |
| Tabla 14. Matriz de Confusión Modelo Original en el periodo 2005 – 2006..... | 67 |
| Tabla 15. Punto de Corte para Clasificación de Clientes durante el periodo 2005 – 2006 | 69 |
| Tabla 16. Punto de Corte según Estrategia de Costo Asociado..... | 70 |