

## TABLA DE CONTENIDO

1.	INTRODUCCIÓN.....	7
1.1.	ANTECEDENTES GENERALES.....	7
1.2.	DEFINICIÓN Y JUSTIFICACIÓN DEL PROYECTO .....	11
1.3.	OBJETIVOS .....	18
1.3.1.	GENERAL.....	18
1.3.2.	ESPECÍFICOS .....	18
1.4.	ALCANCES .....	18
1.5.	RESULTADOS ESPERADOS.....	19
2.	MARCO CONCEPTUAL .....	20
2.1.	DEFINICIÓN DE FRAUDE .....	20
2.2.	TIPOS DE FRAUDE .....	20
2.3.	DEFINICIÓN DE FRAUDE INTERNO.....	22
3.	MARCO TEÓRICO .....	23
3.1.	EL PROCESO KDD .....	23
3.1.1.	RECOLECCIÓN Y PREPARACIÓN DE LOS DATOS .....	25
3.1.2.	PRE-PROCESAMIENTO DE LOS DATOS.....	25
3.1.3.	TRANSFORMACIÓN DE LOS DATOS .....	26
3.1.4.	DATA MINING .....	27
3.1.4.1.	MODELOS SUPERVISADOS.....	28
3.1.4.2.	MODELOS NO SUPERVISADOS .....	29
3.1.4.2.1.	ANÁLISIS DE VECINOS MÁS CERCANOS .....	30
3.1.4.2.2.	ANÁLISIS DE COMPONENTES PRINCIPALES (PCA) .....	31
3.1.4.2.3.	LOCAL OUTLIER FACTOR (LOF) .....	32
3.1.5.	ANÁLISIS E INTERPRETACIÓN DE RESULTADOS .....	33
3.2.	HERRAMIENTAS TECNOLÓGICAS .....	34
4.	APLICACIÓN DE TÉCNICAS DE DATA MINING .....	35
4.1.	RECOLECCIÓN Y PREPARACIÓN DE LOS DATOS .....	35
4.1.1.	CARACTERIZACIÓN Y ESTRUCTURA DE LOS DATOS .....	35
4.1.2.	CARACTERIZACIÓN DE LOS REGISTROS DE FRAUDE .....	36
4.1.3.	SELECCIÓN DE LOS DATOS .....	38

4.2.	PRE-PROCESAMIENTO DE LOS DATOS.....	39
4.2.1.	LIMPIEZA DE LOS DATOS .....	39
4.3.	TRANSFORMACIÓN DE LOS DATOS .....	40
4.3.1.	GENERACIÓN Y SELECCIÓN DE VARIABLES .....	40
4.3.2.	ANÁLISIS PRELIMINAR DE LAS VARIABLES .....	42
4.3.3.	NORMALIZACIÓN DE VARIABLES .....	50
4.4.	DATA MINING .....	50
4.4.1.	ANÁLISIS DE VECINOS MÁS CERCANOS .....	51
4.4.2.	LOCAL OUTLIER FACTOR (LOF) .....	55
4.4.3.	ANÁLISIS DE COMPONENTES PRINCIPALES (PCA) .....	61
4.5.	ANÁLISIS E INTERPRETACIÓN DE RESULTADOS .....	62
5.	CONCLUSIONES Y TRABAJOS FUTUROS .....	66
6.	BIBLIOGRAFÍA.....	68
7.	ANEXOS.....	70
	ANEXO A: EVOLUCIÓN DE INDICADORES DE PROFUNDIDAD FINANCIERA Y COMPARACIÓN CON OTROS PAÍSES.....	70
	ANEXO B: CORRELACIÓN ENTRE INGRESO PER-CÁPITA Y PROFUNDIDAD BANCARIA .....	71
	ANEXO C: EVOLUCIÓN DEL INGRESO PER-CÁPITA EN CHILE.....	72
	ANEXO D: EVOLUCIÓN DE LA PENETRACIÓN DE TARJETAS DE CRÉDITO, DÉBITO Y CUENTAS CORRIENTES.....	73
	ANEXO E: PENETRACIÓN DE LAS TARJETAS DE CRÉDITO POR PAÍS .....	74
	ANEXO F: PAÍSES EN QUE EVERIS TIENE PRESENCIA.....	75
	ANEXO G: FACTURACIÓN POR OFICINA .....	76
	ANEXO H: FACTURACIÓN POR ACTIVIDAD .....	77
	ANEXO I: ARQUITECTURA BIG DATA DE EVERIS .....	78
	ANEXO J: FUNCIONES DE LA ARQUITECTURA BIG DATA EVERIS.....	79
	ANEXO K: MUESTRA DE LA BASE DEL LOG DE TECLEO .....	80
	ANEXO L.1: 30 VECINOS MÁS CERCANOS AL FRAUDE CCON818.....	81
	ANEXO L.2: 30 VECINOS MÁS CERCANOS AL FRAUDE JTAPIAL .....	82
	ANEXO L.3: 30 VECINOS MÁS CERCANOS AL FRAUDE CCON243 (1) .....	83
	ANEXO L.4: 30 VECINOS MÁS CERCANOS AL FRAUDE CCON643.....	84
	ANEXO L.5: 30 VECINOS MÁS CERCANOS FRAUDE CCON243 .....	85

ANEXO M.1: DISTANCIA DE VECINOS MÁS CERCANOS A CCON818 A MEDIA DE FRAUDES .....	86
ANEXO M.2: DISTANCIA DE VECINOS MÁS CERCANOS A JTAPIAL A MEDIA DE FRAUDES .....	87
ANEXO M.3: DISTANCIA DE VECINOS MÁS CERCANOS A CCON243 (1) A MEDIA LOS FRAUDES .....	88
ANEXO M.4: DISTANCIA DE VECINOS MÁS CERCANOS A CCON643 A MEDIA DE FRAUDES .....	89
ANEXO M.5: DISTANCIA DE VECINOS MÁS CERCANOS A CCON243 A MEDIA DE FRAUDES .....	90
ANEXO N: USUARIOS CLASIFICADOS COMO OUTLIERS POR LOF .....	91
ANEXO O: DISTANCIA DE LOS SOSPECHOSOS LOF A LA MEDIA DE LOS DATOS .....	92
ANEXO P: RESULTADO ANÁLISIS DE LAS COMPONENTES PRINCIPALES .	93
ANEXO Q: LISTA COMPLETA DE SOSPECHOSOS ENCONTRADOS POR LOS TRES MODELOS .....	98
ANEXO S: DISTANCIA DE SOSPECHOSOS DE LA LISTA FINAL A LA MEDIA DE LOS DATOS .....	104