

## Índice de contenidos

---

Objetivos .....	0
Índice de contenidos.....	1
Índice de figuras.....	3
Índice de tablas .....	4
1. Introducción.....	5
1.1. Conceptos .....	5
1.1.1. Knowledge Discovery in Databases.....	5
1.1.2. Data Mining .....	6
1.1.3. CRM .....	6
1.1.4. XML.....	7
1.1.5. Ruby.....	7
1.1.6. Ruby on Rails (RoR) .....	7
1.1.7. Arquitectura Modelo-Vista-Controlador.....	8
1.1.8. RubyGems .....	8
1.1.9. Subversion .....	8
1.2. Hardware .....	9
1.3. Software .....	9
1.3.1. Sistemas operativos .....	9
1.3.2. Virtualización .....	9
1.3.3. Entornos de desarrollo .....	9
1.3.4. Frameworks.....	10
1.3.5. Gestor de versiones.....	10
1.4. Fases de diseño y desarrollo.....	10
1.4.1. Primera fase: Prototipo .....	10
1.4.2. Segunda fase: Relación con el CRM .....	11
1.4.3. Fase final: Implementación final .....	12
1.5. Estructura del documento .....	13
2. Fundamentos de Minería de Datos .....	15
2.1. Data Mining y el proceso KDD .....	15
2.2. Procedimientos de Data Mining .....	16
2.3. Algoritmos en procedimientos de Data Mining. Estado del Arte. ....	17
2.3.1. Clustering .....	17
2.3.2. Clasificación.....	21

2.3.3.	Reglas de Asociación .....	24
3.	Diseño y desarrollo .....	27
3.1.	Primera fase: Prototipo .....	28
3.1.1.	Estudio metódico de los algoritmos de Data Mining .....	28
3.1.2.	Selección de algoritmos .....	28
3.1.3.	Diseño de un prototipo .....	30
3.1.4.	Implementación del algoritmo kmedoids .....	40
3.1.5.	Construcción de una base de datos de prueba.....	41
3.2.	Segunda fase: Relación con el CRM.....	41
3.2.1.	Adaptación del prototipo .....	41
3.2.2.	Diseño de los campos de la base de datos.....	44
3.2.3.	Adaptación al algoritmo elegido para los casos de clientes y ofertas 47	
3.2.4.	Almacenamiento de la salida del algoritmo de clustering.....	52
3.2.5.	Empaquetamiento del algoritmo de minería de datos en “gem” ....	54
3.2.6.	Implementación de la interfaz de acceso a clientes, ofertas y operaciones (CRM) .....	54
3.3.	Fase final: Implementación final .....	57
3.3.1.	Pasos para la construcción del programa .....	57
3.3.2.	Algoritmo de reglas .....	58
3.3.3.	Diseño del modelo de datos final de la aplicación.....	63
3.3.4.	Implementación del modelo final .....	72
4.	Descripción de la interfaz realizada .....	75
4.1.	Requisitos del sistema .....	75
4.2.	Diagrama de flujo del operador .....	76
4.3.	Operación de la interfaz .....	79
4.3.1.	Página principal y menú .....	79
4.3.2.	Entidades .....	80
4.3.3.	Campos .....	82
4.3.4.	Valores de discretización.....	85
4.3.5.	Procesado del XML.....	88
4.3.6.	Búsqueda de reglas .....	89
4.3.7.	Visualización de reglas .....	90
5.	Conclusiones y Líneas Futuras .....	93
	Glosario .....	95
	Referencias.....	97
	Anexo: Manual de usuario	

## Índice de figuras

---

Figura 1. Descomposición del proceso KDD .....	15
Figura 2. Diagrama de bloques general.....	31
Figura 3. Descomposición de la sección correspondiente a procedimientos aplicados a clientes.....	31
Figura 4. Descomposición de la sección correspondiente a procedimientos aplicados a ofertas.....	32
Figura 5. Detalle de la sección correspondiente a acciones aplicadas a oportunidades. ....	34
Figura 6. Detalle de la sección correspondiente a procedimientos aplicados para la obtención de los resultados finales.....	35
Figura 7. Diseño del módulo de Data Mining para una aplicación de comercio electrónico.....	40
Figura 8. Diseño intermedio del algoritmo de Data Mining aplicado a comercio electrónico.....	42
Figura 9. Diseño intermedio expandido del algoritmo de Data Mining aplicado a comercio electrónico.....	43
Figura 10. Estructura del código postal.....	48
Figura 11. Uso habitual del algoritmo Apriori.....	59
Figura 12. Uso adaptado del algoritmo Apriori.....	60
Figura 13. Formato del fichero XML.....	62
Figura 14. Modelo final de datos de aplicación.....	68
Figura 15. Diagrama de flujo del operador.....	77
Figura 16. Página principal.....	79
Figura 17. Listar entidades.....	81
Figura 18. Añadir entidad.....	81
Figura 19. Editar entidad.....	82
Figura 20. Listar campos.....	83
Figura 21. Añadir campos.....	83
Figura 22. Editar campos.....	84
Figura 23. Listar valores de discretización.....	85
Figura 24. Añadir valor de discretización.....	86
Figura 25. Editar valor de discretización.....	87
Figura 26. Selección del archivo XML.....	88
Figura 27. Procesado del XML concluido.....	88
Figura 28. Búsqueda de reglas.....	89
Figura 29. Búsqueda de reglas concluida.....	90
Figura 30. Reglas sin filtrar.....	91
Figura 31. Reglas filtradas.....	92

## Índice de tablas

---

Tabla 1. Campos y valores restringidos de clientes.....	47
Tabla 2. Diez ciudades más pobladas según datos del INE, y sus correspondientes códigos postales. ....	49
Tabla 3. Campos y valores restringidos de ofertas.....	50
Tabla 4. Tabla llamada “minería_pre_clusterings” . ....	52
Tabla 5. Tabla “minería_res_clusterings”.....	53
Tabla 6. Almacenamiento previo al clustering de ofertas.....	53
Tabla 7. Almacenamiento de resultados del clustering de ofertas. ....	53
Tabla 8. Campos y valores restringidos de clientes.....	66
Tabla 9. Tabla de reglas prototipo.....	66
Tabla 10. Tabla de antecedentes prototipo. ....	67
Tabla 11. Tabla de consecuentes prototipo. ....	67
Tabla 12. Tabla de Entidades.....	68
Tabla 13. Tabla de Métodos. ....	69
Tabla 14. Tabla de Campos.....	69
Tabla 15. Tabla de Valores de Discretización. ....	70
Tabla 16. Tabla de Características. ....	70
Tabla 17. Tabla de Transacciones.....	70
Tabla 18. Tabla de Relaciones entre Transacciones y Características.....	71
Tabla 19. Tabla de Reglas. ....	71
Tabla 20. Tabla de Antecedentes. ....	71
Tabla 21. Tabla de Consecuentes. ....	72