

ÍNDICE SISTEMÁTICO

	PÁGINA
Sumario	5
Prólogo	7
Unidad didáctica 1. Introducción a la minería de datos	9
Objetivos de la Unidad	10
1. Origen y motivación	11
2. El proceso de descubrimiento de conocimiento en bases de datos	13
3. Tipos de datos	15
4. Relación con otras áreas	19
5. Utilidad de la minería de datos	20
5.1. Ejemplo 1: grupos de clientes en una compañía telefónica	20
5.2. Ejemplo 2: estudio de la concesión de créditos en un banco	21
5.3. Ejemplo 3: análisis de la cesta de la compra en un supermercado	22
6. Aplicaciones	24
7. Repercusión, desafíos y tendencias en minería de datos	25
8. Conclusiones	28

Conceptos básicos a retener	29
Actividades de autocomprobación	30
Actividades de repaso	33
Ejercicios voluntarios	34
Referencias bibliográficas	35
Unidad didáctica 2. El proceso de KDD.....	37
Objetivos de la Unidad	38
1. Introducción	39
2. Recopilación de datos	42
3. Selección, limpieza y transformación de datos	44
3.1. Selección de datos	44
3.2. Limpieza de datos	45
3.3. Transformación de datos	46
4. Data mining	49
5. Interpretación y evaluación	51
6. Visión de conjunto	53
7. Conclusiones	54
Conceptos básicos a retener	56
Actividades de autocomprobación	57
Actividades de repaso	60
Ejercicios voluntarios	62
Referencias bibliográficas	63
Unidad didáctica 3. Modelo de data mining	65
Objetivos de la Unidad	66
1. Introducción	67
2. Clasificación	68

3. Regresión	70
4. Clustering	71
5. Asociación	72
6. Detección de atípicos	73
7. Tareas y técnicas	75
8. Técnicas y modelos	76
9. Herramientas de data mining	78
10. Conclusiones	81
Conceptos básicos a retener	83
Actividades de auto comprobación	84
Actividades de repaso	88
Ejercicios voluntarios	89
Referencias bibliográficas	91
Unidad didáctica 4. Clasificación: árboles de decisión y redes de neuronas artificiales	93
Objetivos de la Unidad	94
1. Introducción	95
2. Árboles y reglas de decisión	97
3. Redes de neuronas artificiales	101
4. Clasificación con Weka	104
4.1. Árboles y reglas de decisión con Weka	106
4.2. Redes de neuronas artificiales con Weka	111
5. Conclusiones	114
Conceptos básicos a retener	116
Actividades de auto comprobación	117
Actividades de repaso	120
Ejercicios voluntarios	121
Referencias bibliográficas	122

Unidad didáctica 5. Clasificación: técnicas bayesianas y técnicas basadas en casos	123
Objetivos de la Unidad	124
1. Introducción	125
2. Técnicas bayesianas	127
3. Técnicas basadas en casos	130
4. Clasificación con Weka	134
4.1. Técnicas bayesianas con Weka	135
4.2. Técnicas basadas en casos con Weka	138
5. Evaluación de clasificadores	141
6. Conclusiones	144
Conceptos básicos a retener	145
Actividades de autocomprobación	145
Actividades de repaso	149
Ejercicios voluntarios	151
Referencias bibliográficas	152
Unidad didáctica 6. Clustering: técnicas particionales	153
Objetivos de la Unidad	154
1. Introducción	155
2. Medidas de distancia	158
3. Clustering particional	160
4. Clustering con Weka	165
4.1. Clustering particional con Weka	167
5. Conclusiones	171
Conceptos básicos a retener	172
Actividades de autocomprobación	173
Actividades de repaso	175
Ejercicios voluntarios	176
Referencias bibliográficas	177

Unidad didáctica 7. Clustering: técnicas jerárquicas y técnicas basadas en densidad	179
Objetivos de la Unidad	180
1. Introducción	181
2. Clustering jerárquico	182
3. Clustering basado en densidad	190
4. Clustering basado en grid	193
5. Clustering con Weka	194
5.1. Clustering jerárquico con Weka	194
5.2. Clustering basado en densidad con Weka	197
6. Evaluación de modelos de clustering	200
7. Conclusiones	202
Conceptos básicos a retener	203
Actividades de auto comprobación	204
Actividades de repaso	207
Ejercicios voluntarios	208
Referencias bibliográficas	209
Unidad didáctica 8. Asociación, regresión y detección de atípicos	211
Objetivos de la Unidad	212
1. Introducción	213
2. Asociación	213
3. Regresión	216
3.1. Regresión lineal	216
3.2. Regresión no lineal	218
4. Detección de atípicos	219
5. Asociación con Weka	222
5.1. Apriori con Weka	222
6. Conclusiones	225

Conceptos básicos a retener	227
Actividades de autocomprobación	227
Actividades de repaso	231
Ejercicios voluntarios	232
Referencias bibliográficas	233
Unidad didáctica 9. Minería de datos no convencionales	235
Objetivos de la Unidad	236
1. Introducción	237
2. Análisis de series temporales	238
2.1. Comparación de series temporales	240
2.2. Análisis de subsecuencias	242
2.3. Modelos de referencia de series temporales	243
2.4. Análisis de eventos en series temporales	244
3. Web mining	245
4. Text mining	247
5. Data mining espacial	249
6. Minería de datos multimedia	250
7. Conclusiones	251
Conceptos básicos a retener	253
Actividades de autocomprobación	254
Actividades de repaso	257
Ejercicios voluntarios	258
Referencias bibliográficas	259
Unidad didáctica 10. Caso de estudio: aplicación de la minería de datos en el dominio de la medicina	261
Objetivos de la Unidad	262

1. Introducción	263
2. Dominio médico de aplicación: estabilometría	263
2.1. Unilateral Stance (UNI)	264
2.2. Rhythmic Weight Shift (RWS)	266
2.3. Limits of Stability (LOS)	267
3. Marco para el descubrimiento de conocimiento en el entorno médico	268
3.1. Definición de eventos en series temporales	269
3.2. Comparación entre dos series temporales	271
3.3. Generación de modelos de referencia de series temporales	274
3.4. Sistema implementado	277
4. Conclusiones	278
Conceptos básicos a retener	280
Actividades de autocomprobación	281
Actividades de repaso	284
Ejercicios voluntarios	285
Referencias bibliográficas	285