

Índice general

Prólogo	ix
CAPÍTULO 1. Fundamentos del Machine Learning	1
1.1. Introducción	2
1.2. Conceptos sobre datos	2
1.3. Conceptos sobre aprendizaje	3
1.4. Tipos de problemas	4
1.5. Tipos de datos	4
1.6. Tipos de aprendizajes	5
1.7. Etapas de implementación del aprendizaje máquina	6
1.8. Exploración y preparación de datos	7
1.9. Visualización de datos	15
Referencias	20
CAPÍTULO 2. Bases matemáticas	21
2.1. Introducción	22
2.2. Probabilidad	22
2.2.1. Variables aleatorias discretas	22
2.2.2. Reglas fundamentales	23
2.2.2.1. Probabilidad de la unión de dos elementos	23
2.2.2.2. Probabilidad de la intersección de dos elementos	23
2.2.2.3. Probabilidad condicional	24
2.2.3. Algunas distribuciones comunes	25
2.2.3.1. Distribución binomial y Bernoulli	25
2.2.3.2. Distribución multinomial	26
2.2.3.3. Distribución de Poisson	27
2.2.3.4. Distribución uniforme	27
2.2.3.5. Distribución normal (gaussiana)	28
2.3. Estadística	29
2.3.1. Medidas de tendencia central	30
2.3.1.1. Media aritmética	31
2.3.1.2. Mediana	31
2.3.1.3. Moda	31
2.3.2. Medidas de variabilidad	34
2.3.2.1. Varianza	35
2.3.2.2. Desviación estándar	35
2.3.2.3. Rango	36
2.3.3. Herramientas gráficas	38
2.3.3.1. Gráfica de líneas	38
2.3.3.2. Gráfica de barras	39

2.3.3.3. Gráfica de cajas.....	42
2.3.3.4. Histograma	44
2.4. Álgebra lineal.....	46
2.4.1. Vectores y matrices.....	46
2.4.2. Suma, resta y producto escalar de vectores	46
2.4.3. Norma vectorial.....	47
2.4.3.1. Norma L^1	47
2.4.3.2. Norma L^2	48
2.4.4. Matrices.....	48
2.4.4.1. Suma, resta y multiplicación de matrices.....	49
2.4.4.2. Multiplicación escalar de una matriz.....	49
2.4.4.3. Multiplicación de matrices	50
2.4.5. Tipo de matrices.....	51
2.4.6. Descomposición de matrices	53
2.4.6.1. Descomposición LU	53
2.4.6.2. Descomposición de los valores y vectores propios	54
Referencias	55
CAPÍTULO 3. Clasificación	57
3.1. Introducción	58
3.2. Vecinos cercanos (k-NN)	58
3.2.1. Ejemplo de clasificación utilizando k-NN en MATLAB.....	61
3.3. Regresión logística.....	62
3.3.1. Ejemplo de clasificación utilizando la regresión logística en MATLAB	63
3.4. Naive Bayes	64
3.4.1. Teorema de Bayes	65
3.4.2. Clasificador Naive Bayes.....	65
3.4.3. Ejemplo de clasificación Naive Bayes utilizando MATLAB	67
3.5. Análisis del discriminante de Fisher	68
3.5.1. Ejemplo de clasificación mediante discriminante de Fisher utilizando MATLAB	71
3.6. Máquina de vector soporte (SVM).....	73
3.6.1. Ejemplo de clasificación con máquina vector soporte utilizando MATLAB	78
Referencias.....	79
CAPÍTULO 4. Regresión lineal.....	81
4.1. Introducción	82
4.2. Regresión lineal simple.....	83
4.3. Mínimos cuadrados.....	85
4.4. Gradiente descendente en regresión lineal simple.....	90

4.5. Ecuación normal en regresión lineal simple.....	95
4.6. Regresión lineal múltiple.....	101
Referencias.....	110
CAPÍTULO 5. Agrupamiento (<i>clustering</i>).....	111
5.1. Introducción.....	112
5.2. Algoritmo de K-means.....	112
5.2.1. El uso de K-means en MATLAB.....	116
5.3. Método de expectación-maximización.....	118
5.3.1. Modelos de mezclas gaussianas.....	118
5.3.2. Estimación de máxima probabilidad.....	119
5.3.3. EM en una dimensión.....	120
5.3.4. EM en varias dimensiones.....	123
5.4. Agrupación jerárquica.....	128
5.4.1. Medidas de similaridad entre los clústeres.....	130
5.5. El algoritmo de Fuzzy C-means.....	134
Referencias.....	137
CAPÍTULO 6. Reducción de dimensionalidad.....	139
6.1. Introducción.....	140
6.2. Análisis de componentes principales (PCA).....	140
6.2.1. Ejemplo en MATLAB de análisis de componentes principales (PCA).....	144
6.3. Análisis de componentes independientes (ICA).....	146
6.3.1. Distribuciones gaussianas no permitidas.....	147
6.3.2. Estimación de máxima probabilidad.....	148
6.3.3. Ejemplo en MATLAB de análisis de componentes independientes (ICA).....	149
6.4. Análisis de factor (FA).....	151
6.4.1. Ejemplo en MATLAB de análisis de factor (FA).....	153
Referencias.....	155
CAPÍTULO 7. Métodos unidos.....	157
7.1. Árboles de decisión.....	158
7.2. Algoritmo CART.....	160
7.3. Árboles de decisión para clasificación.....	160
7.4. Árboles de decisión para regresión.....	163
7.5. Funciones de MATLAB para árboles de decisión.....	164
7.6. El método Bootstrap.....	170
Referencias.....	174
CAPÍTULO 8. Reconocimiento de objetos.....	175
8.1. Comparación en imágenes a escala de grises.....	176

8.2. Distancia entre patrones	177
8.3. Distancia y correlación	181
8.4. La correlación cruzada normalizada.....	184
8.5. Coeficiente de correlación	186
8.6. Reconocimiento de patrones usando el coeficiente de correlación	190
8.7. Comparación de imágenes binarias	196
8.7.1. La transformación de distancia	196
8.7.2. El algoritmo de Chamfer	198
8.8. Índice de relación de Chamfer	202
Referencias.....	206
CAPÍTULO 9. Estadística inferencial.....	207
9.1. Introducción	208
9.2. Distribución de muestreo.....	209
9.2.1. Distribución normal.....	210
9.2.2. Distribución t.....	216
9.3. Estimación de parámetros	221
9.3.1. Estimación por intervalos.....	223
9.4. Pruebas de hipótesis	227
Referencias.....	231
CAPÍTULO 10. Evaluación del desempeño	233
10.1. Introducción	233
10.2. Ajustes sobre las predicciones de un clasificador	234
10.2.1. Tipos de errores	236
10.2.2. Matriz de confusión	236
10.3. Métricas para clasificadores.....	236
10.3.1. Exactitud (ACC).....	237
10.3.2. Razón de falsos positivos (FPR)	237
10.3.3. Sensibilidad (VPR).....	237
10.3.4. Especificidad (SPC)	237
10.3.5. Precisión	237
10.3.6. F1.....	238
10.3.7. F2.....	238
10.4. Curva ROC.....	238
10.5. El balance entre el sesgo y la varianza	239
10.6. Evaluación de modelos.....	240
10.6.1. Técnicas comunes de validación cruzada.....	241
10.7. Métricas de error en regresores lineales	243
Referencias.....	245